



XX PD 25-OCT-2001.  
 XX XX  
 XX PF 16-APR-2001; 2001WO-US08656.  
 XX XX  
 XX PR 18-APR-2000; 2000US-0552929.  
 XX PR 26-JAN-2001; 2001US-0770160.  
 XX (HYSE-) HYSEQ INC.  
 XX PA  
 XX XX  
 XX PI Tang YT, Liu C, Drmanac RT;  
 XX XX  
 XX DR WPI; 2001-611725/70.  
 XX XX  
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
 XX PT vaccination, testing and therapy  
 XX XX  
 XX PS Claim 20; Page 613; 765pp; English.  
 XX XX  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising  
 CC the nucleic acids encoding the polypeptides and cells genetically  
 CC engineered to express them are also useful for producing the proteins.  
 CC The proteins are useful in genetic vaccination, testing and  
 CC therapy, and can be used as nutritional supplements. They may be used to  
 CC increase stem cell proliferation; to regulate haematopoiesis; and in  
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.  
 XX XX  
 SQ Sequence 648 AA;  
 Query Match 46.8%; Score 288; DB 22; Length 648;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-282;  
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSGIKKQKTENQKSTNNVYQAAHVSRNKRQGVVGRGCTVWLTGLSGAGKTIISF 60  
 DB 27 MSGIKKQKTENQKSTNNVYQAAHVSRNKRQGVVGRGCTVWLTGLSGAGKTIISF 86  
 QY 61 ALEEYLVSHPAICYSLDGDNVRHGLNRLNLFSPGDRNIRIRIAEVAKLFDADAGLCVITS 120  
 DB 87 ALEEYLVSHPAICYSLDGDNVRHGLNRLNLFSPGDRNIRIRIAEVAKLFDADAGLCVITS 146  
 QY 121 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGEIKGFTGIDS 180  
 DB 147 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGEIKGFTGIDS 206  
 QY 181 DYEKPTPEPVLKTNLSVSDCVHQQVVELLQEQNIIVPYTIIRKDIHELFPENKLDHVRAE 240  
 DB 207 DYEKPTPEPVLKTNLSVSDCVHQQVVELLQEQNIIVPYTIIRKDIHELFPENKLDHVRAE 266  
 QY 241 AETPLSLSTIKLDLQWVQLSEGWATPLKGFMRKEYLQVMHFDTLDD 288  
 DB 267 AETPLSLSTIKLDLQWVQLSEGWATPLKGFMRKEYLQVMHFDTLDD 314

RESULT 2

AAV79214  
 ID AAY79214 standard; Protein; 619 AA.  
 XX XX  
 XX AC AAY79214;  
 XX XX  
 XX DT 19-JUN-2000 (first entry)  
 XX XX  
 XX DT Human transferase TRNSFS-6.  
 XX XX

XX Transferase; TRNSFS-6; human; antitumour; cancer;  
 KW gastrointestinal disorder; developmental disorder;  
 KW genetic disorder; neurological disorder; reproductive disorder;  
 KW smooth muscle disorder; immunological disorder; inflammation;  
 KW diagnosis; therapy; ATP sulfurylase/APS kinase 2.  
 XX XX

OS Homo sapiens.

XX Key Location/Qualifiers  
 XX Modified-site 92  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 176  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 180  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 198  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 250  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 285  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 308  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 313  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 394  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 453  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 527  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 548  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 430  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 439  
 FT /note= "potential O-phosphorylation"  
 FT Modified-site 195  
 FT /note= "potential N-glycosylation"  
 FT Modified-site 298  
 FT /note= "potential N-glycosylation"  
 FT Binding-site 49..56  
 FT /note= "ATP/GTP binding site (P-loop)"  
 XX WO200014251-A2.

XX 16-MAR-2000.  
 XX 09-SEP-1999; 99WO-US20989.  
 XX 10-SEP-1998; 98US-0150657.  
 PR 04-NOV-1998; 98US-0186779.  
 PR 11-MAY-1999; 99US-0133642.  
 XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;  
 PI Hillman JL, Azimzai Y;  
 XX WPI; 2000-256996/22.  
 DR N-PSDB; AA294206.

XX Human transferase proteins useful for preventing, diagnosing and  
 PT treating cancers and developmental, gastrointestinal, genetic,  
 PT immunological, neurological, reproductive and smooth muscle disorders -  
 XX Claim 1; Page 84-85; 113pp; English.

XX The present sequence is that of human transferase TRNSFS-6, 1 of  
 CC 15 claimed novel human transferase proteins of the invention (see  
 CC AAY79209-23). The sequence was deduced from a cDNA clone (see  
 CC AA294206) isolated from a foetal kidney library. It shows homology

CC to human ATP sulfurylase/APS kinase 2. TRNSFS-6 is expressed in  
 CC cardiovascular, gastrointestinal, reproductive and endocrine  
 CC tissues, especially those associated with cancer, inflammation and  
 CC cell proliferation. The new human transferases and polynucleotides  
 CC can be used in the diagnosis, prevention and treatment of cancer,  
 CC developmental disorders, gastrointestinal disorders, genetic  
 CC disorders, immunological disorders, neurological disorders,  
 CC reproductive disorders, and smooth muscle disorders. The  
 CC polypeptides can also be used to raise antibodies, and to screen  
 CC for agonists and antagonists of transferase activity.

XX Sequence 619 AA;  
 SQ Query Match 35.9%; Score: 221; DB 21; Length 619;  
 Best Local Similarity 99.7%; Pred. No. 3.7e-214;  
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 289 DGVINNSIPIVLPVSAEDKTRLECCSKFVLAHGRVAILRDAAEFYEHKKEECRSRWGT 348  
 DB 294 DGVINNSIPIVLPVSAEDKTRLECCSKFVLAHGRVAILRDAAEFYEHKKEECRSRWGT 353  
 QY 349 TCTKHPHKIMVMSGDLVGGDLQVLEKIRWNDGLDQYRLTPLELKOKCKEMNADAVFAF 408  
 DB 354 TCTKHPHKIMVMSGDLVGGDLQVLEKIRWNDGLDQYRLTPLELKOKCKEMNADAVFAF 413  
 QY 409 QLRNPVNHGHALLMQDTCRLLRGYKHPVLLHPLGGWTKDDVPLDMRMKHAVALLEE 468  
 DB 414 QLRNPVNHGHALLMQDTCRLLRGYKHPVLLHPLGGWTKDDVPLDMRMKHAVALLEE 473  
 QY 469 GVLDPKSTIVAIFPSPLVAGPTEVQWHCSRMIAAGANFYIVGRDPAGMPHETKDKLYE 528  
 DB 474 GVLDPKSTIVAIFPSPLVAGPTEVQWHCSRMIAAGANFYIVGRDPAGMPHETKDKLYE 533  
 QY 529 PTHGGKVLSPAGLTSVEIIPFVAAYNKAKKAMDFYDPAHNEFFDFISGTRMKLAREG 588  
 DB 534 PTHGGKVLSPAGLTSVEIIPFVAAYNKAKKAMDFYDPAHNEFFDFISGTRMKLAREG 593  
 QY 589 ENPPDGFMAPKANKVLTDIYRS 610  
 DB 594 ENPPDGFMAPKANKVLTDIYRS 615

RESULT 3  
 ABB40266  
 ID ABB40266 standard; Peptide: 65 AA.  
 XX AC ABB40266;  
 XX DT 04-FEB-2002 (first entry)  
 XX DE Peptide #7772 encoded by human foetal liver single exon probe.  
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX OS Homo sapiens.  
 XX PN WO200157277-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US00669.  
 XX PR 04-FEB-2000; 2000US-0180312.  
 XX PR 26-MAY-2000; 2000US-0207456.  
 XX PR 30-JUN-2000; 2000US-0608408.  
 XX PR 03-AUG-2000; 2000US-0632366.  
 XX PR 21-SEP-2000; 2000US-0234687.  
 XX PR 27-SEP-2000; 2000US-0236359.  
 XX PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-488899/53.

XX WPI; 2001-483447/52.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human foetal liver.  
 XX PT Claim 27; SEQ ID NO 32901; 639pp + sequence listing; English.  
 XX CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC foetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 65 AA;  
 SQ Query Match 10.6%; Score 65; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 63 EEVLVSHAIPCYSLDGDGNVRHGLNRNLGFSFGDRENNIRIAEVAKLFPADAGLVCTISFI 122  
 DB 1 EEVLVSHAIPCYSLDGDGNVRHGLNRNLGFSFGDRENNIRIAEVAKLFPADAGLVCTISFI 60  
 QY 123 SPEAK 127  
 DB 61 SPEAK 65

RESULT 4  
 ABB24683  
 ID ABB24683 standard; Protein: 65 AA.  
 XX AC ABB24683;  
 XX DT 23-JAN-2002 (first entry)  
 XX DE Protein #682 encoded by probe for measuring heart cell gene expression.  
 XX KW Human; gene expression; heart; microarray; vascular system;  
 XX KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 XX KW congenital heart disease.  
 XX OS Homo sapiens.  
 XX PN WO200157274-A2.  
 XX PD 09-AUG-2001.  
 XX PF 30-JAN-2001; 2001WO-US00666.  
 XX PR 04-FEB-2000; 2000US-0180312.  
 XX PR 26-MAY-2000; 2000US-0207456.  
 XX PR 30-JUN-2000; 2000US-0608408.  
 XX PR 03-AUG-2000; 2000US-0632366.  
 XX PR 21-SEP-2000; 2000US-0234687.  
 XX PR 27-SEP-2000; 2000US-0236359.  
 XX PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX PI WPI; 2001-488899/53.  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts -  
 XX Claim 15; SEQ ID NO 26453; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 65 AA;

Query Match 10.6%; Score 65; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDRNIRRIAEVAKLFADAGLVCTISFI 122  
 DB 1 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDRNIRRIAEVAKLFADAGLVCTISFI 60  
 QY 123 SPFAK 127  
 DB 61 SPFAK 65

RESULT 5  
 AAM61056  
 ID AAM61056 standard; Protein; 65 AA.  
 AC AAM61056;  
 XX  
 XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33161.  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX Homo sapiens.

XX WO200157275-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human  
 XX brains -  
 XX Example 4; SEQ ID NO: 33161; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 65 AA;

Query Match 10.6%; Score 65; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDRNIRRIAEVAKLFADAGLVCTISFI 122  
 DB 1 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDRNIRRIAEVAKLFADAGLVCTISFI 60  
 QY 123 SPFAK 127  
 DB 61 SPFAK 65

RESULT 6  
 AAM73760  
 ID AAM73760 standard; Protein; 65 AA.  
 AC AAM73760;  
 XX  
 XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 34066.  
 XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX Homo sapiens.

XX WO200157276-A2.  
 XX 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.  
 XX 26-MAY-2000; 2000US-0207456.  
 XX 30-JUN-2000; 2000US-0608408.  
 XX 03-AUG-2000; 2000US-0632366.  
 XX 21-SEP-2000; 2000US-0234687.  
 XX 27-SEP-2000; 2000US-0236359.  
 XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human bone marrow -  
 XX Example 4; SEQ ID NO: 34066; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX

SQ Sequence 65 AA;  
 Query Match 10.6%; Score 65; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy 63 EYLVSHPAICYSLDGDNVRHGLNRNLGFGSPGDRREINIRIAEVAKLFADAGLVCIITSFI 122  
 Db 1 EYLVSHPAICYSLDGDNVRHGLNRNLGFGSPGDRREINIRIAEVAKLFADAGLVCIITSFI 60  
 Qy 123 SPFAK 127  
 Db 61 SPFAK 65

RESULT 7  
 AAM33947  
 ID AAM33947 standard; Protein: 65 AA.  
 XX  
 AC AAM33947;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #7984 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon-nucleic acid probes useful for  
 PT analyzing gene expression in human placenta

XX Claim 27; SEQ ID No 34216; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AA13315-AA157346). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 XX  
 XX Sequence 65 AA;

Query Match 10.6%; Score 65; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 EYLVSHPAICYSLDGDNVRHGLNRNLGFGSPGDRREINIRIAEVAKLFADAGLVCIITSFI 122  
 Db 1 EYLVSHPAICYSLDGDNVRHGLNRNLGFGSPGDRREINIRIAEVAKLFADAGLVCIITSFI 60  
 Qy 123 SPFAK 127  
 Db 61 SPFAK 65

RESULT 8  
 ABG43644

ID ABG43644 standard; Peptide: 65 AA.  
 XX  
 AC ABG43644;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 33309.  
 XX  
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.  
 XX  
 PF 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to  
 XX measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 33309; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon  
 XX nucleic acid probes for measuring gene expression in a sample derived  
 XX from human lung comprising single exon nucleic acid probes having one of  
 XX 12614 nucleic acid sequences mentioned in the specification, or their  
 XX complements or the 12387 open reading frames derived from the 12614  
 XX probes. Also included are a microarray comprising the novel set of  
 XX probes; the novel set of probes which hybridise at high stringency to a  
 XX nucleic acid expressed in the human lung; measuring gene expression in a  
 XX sample derived from human lung, comprising (a) contacting the array with  
 XX a collection of detectably labeled nucleic acids derived from human lung  
 XX mRNA, and (b) measuring the label detectably bound to each probe of  
 XX the array; identifying exons in a eukaryotic genome, comprising  
 XX (a) algorithmically predicting at least one exon from genomic sequences  
 XX of the eukaryote; and (b) detecting specific hybridisation of detectably  
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 XX having a fragment identical to the predicted exon, the probe is included  
 XX in the above mentioned microarray; assigning exons to a single gene,  
 XX comprising (a) identifying exons from genomic sequence by the method  
 XX above and (b) measuring the expression of each of the exons in several  
 XX tissues and/or cell types using hybridisation to a single exon  
 XX microarrays having a probe with the exon, where a common pattern of  
 XX expression of the exons in the tissues and/or cell types indicates that  
 XX the exons should be assigned to a single gene; a peptide comprising one  
 XX of 12011 sequences, mentioned in the specification, or encoded by the  
 XX probes/open reading frames (ORF). The probes are used for gene  
 XX expression analysis, and for identifying exons in a gene, particularly  
 XX using human lung derived mRNA and for the study of lung diseases  
 XX such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemolysis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension,  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 65 AA;

Query Match 10.6%; Score 65; DB 23; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHPICSLDGNVRHGLNRNLGSPGDRNIRIAEVAKLFDAGLVCTISFI 122  
 Db 1 EYLVSHPICSLDGNVRHGLNRNLGSPGDRNIRIAEVAKLFDAGLVCTISFI 60

QY 123 SPEAK 127

Db 61 SPEAK 65

RESULT 9

ABP41780

ID ABP41780 standard; Protein; 133 AA.

XX AC ABP41780;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HMEJA45, SEQ ID NO:2912.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-0518569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54857.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases -

XX PS Claim 11; SEQ ID No 2912; 2922pp; English.

XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 133 AA;

Query Match 8.9%; Score 55; DB 23; Length 133;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-47;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGIKKQKTENQOKSTNVVYQAHVSRNKRGVGVTRGGFRGCTVMTGLSGAGK 55

Db 56 MSGIKKQKTENQOKSTNVVYQAHVSRNKRGVGVTRGGFRGCTVMTGLSGAGK 110

RESULT 10

AAW70494

ID AAW70494 standard; Protein; 624 AA.

XX AC AAW70494;

XX DT 29-DEC-1998 (first entry)

XX DE Human disease related nucleotide kinase-2 (DRNK-2) protein sequence.

XX KW Human disease related nucleotide kinase-2; DRNK-2; deoxyguanosine kinase;  
 KW p21ras; cell proliferation; oncogenesis; cancer; PAPS;  
 KW immune disorder; neurological dysfunction; gene therapy;  
 KW adenosine 3'-phosphate 5'-phosphosulfate-dependent enzyme motif;  
 KW ATP pyrophosphatase PP-motif.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 27

FT Domain /note= "N-glycosylated"

FT Region 59..86

FT Region /note= "ATP/GTP binding site (p-loop)"

FT Region 176..187

FT Modified-site 303 /note= "adenosine 3'-phosphate 5'-phosphosulfate  
 (PAPS)-dependent enzyme motif"

FT Region /note= "N-glycosylated"

FT Region 411..433

FT Region /note= "ATP pyrophosphatase PP-motif"

XX US5817482-A.

XX PD 06-OCT-1998.  
 XX XX  
 XX PF 20-JUN-1997; 97US-0879561.  
 XX PR 20-JUN-1997; 97US-0879561.  
 XX PA (INCY-) INCYTE PHARM INC.  
 XX PI Bandman O, Corley NC, Guegler KJ, Hawkins PR, Hillman JL;  
 XX DR WPI: 1998-556388/47.  
 XX DR N-PSDB: AAV33482.  
 XX PT Nucleic acids encoding deoxyguanosine kinase - useful for  
 PT recombinant production of the enzyme for treating diseases caused by  
 PT lack of the enzyme e.g. cancers caused through loss of enzyme  
 PT function  
 XX XX  
 XX PS Examples; Fig 2A-2G; 53pp; English.  
 XX CC The present sequence represents a human disease related nucleotide  
 CC kinase-2 (DRNK-2) protein sequence. The DNKR-2 encoding DNA sequence  
 CC was first identified in Incyte Clone 373887 from the lung tissue cDNA  
 CC library (LUNGNOT 02). The DNKR-2 DNA sequence is useful for the  
 CC production of the corresponding recombinant enzyme. The invention  
 CC provides DRNK enzymes, which are deoxyguanosine kinases, which catalyse  
 CC the transfer of a terminal phosphate from adenosine triphosphate (ATP)  
 CC or guanine triphosphate (GTP) to guanosine or guanidine in the regulation  
 CC of cellular levels of GTP and its corresponding nucleoside triphosphate.  
 CC As GTP levels are known to control the activity of certain oncogenic  
 CC proteins e.g. p21ras, a protein involved in cell proliferation and  
 CC oncogenesis, suppression of the enzyme activity causes high ratios of  
 CC GTP:GDP, promoting oncogenesis. Therefore, diseases (e.g. cancers,  
 CC immune disorders and neurological dysfunction) related to this lack of  
 CC activity may be prevented or treated with the recombinant enzyme, or by  
 CC gene therapy based strategies. Anti-sense constructs of the DNKR  
 CC encoding nucleic acids may also be used for inhibition of over-expression  
 CC of the enzyme.  
 XX XX  
 XX SQ Sequence 624 AA;  
 Query Match 6.2%; Score 38; DB 19; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 3e-29;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 YQAHVSRNKRQGVVTRGGFRGCTVMTGLSGAGKTT 57  
 DB 30 YQAHVSRNKRQGVVTRGGFRGCTVMTGLSGAGKTT 67  
 RESULT 11  
 AAY22349  
 ID AAY22349 standard; Protein; 624 AA.  
 XX XX  
 XX AC AAY22349;  
 XX DT 24-SEP-1999 (first entry)  
 XX DE Human APS kinase/ATP sulphurylase protein.  
 XX KW APS kinase/ATP sulphurylase; human; PAPS production;  
 KW 3'-phosphoadenosine 5'-phosphosulphate.  
 XX OS Homo sapiens.  
 XX PN JP11187883-A.  
 XX PD 13-JUL-1999.  
 XX XX  
 XX PF 26-DEC-1997; 97JP-0360387.  
 XX XX  
 XX PD 26-DEC-1997; 97JP-0360387.

XX PA (NIRA) UNITIKA LTD.  
 PA (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.  
 XX XX  
 XX DR WPI: 1999-451549/38.  
 XX DR N-PSDB: AAX84897.  
 XX PT New human-derived APS kinase/ATP sulfurylase gene - useful for  
 PT large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate  
 PT (PAPS)  
 XX PS Claim 1: Page 6-7; 9pp; Japanese.  
 XX XX  
 XX CC This sequence is the human-derived APS kinase/ATP sulphurylase of  
 CC the invention. The enzyme may be used to prepare 3'-phosphoadenosine  
 CC 5'-phosphosulphate (PAPS) on a large scale.  
 XX XX  
 XX SQ Sequence 624 AA;  
 Query Match 6.2%; Score 38; DB 20; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 3e-29;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 20 YQAHVSRNKRQGVVTRGGFRGCTVMTGLSGAGKTT 57  
 DB 30 YQAHVSRNKRQGVVTRGGFRGCTVMTGLSGAGKTT 67  
 RESULT 12  
 AAW67882  
 ID AAW67882 standard; Protein; 625 AA.  
 XX XX  
 XX AC AAW67882;  
 XX DT 25-MAR-1999 (first entry)  
 XX DE Human secreted protein encoded by gene 76 clone HOSFD58.  
 XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 625 /label= unknown  
 FT  
 XX PN WO9842738-A1.  
 XX PD 01-OCT-1998.  
 XX PF 19-MAR-1998; 98WO-US05311.  
 XX PR 30-MAY-1997; 97US-0050937.  
 PR 21-MAR-1997; 97US-0041276.  
 PR 21-MAR-1997; 97US-0041277.  
 PR 21-MAR-1997; 97US-0041281.  
 PR 21-MAR-1997; 97US-0042344.  
 PR 30-MAY-1997; 97US-0048069.  
 PR 30-MAY-1997; 97US-0048094.  
 PR 30-MAY-1997; 97US-0048095.  
 PR 30-MAY-1997; 97US-0048096.  
 PR 30-MAY-1997; 97US-0048099.  
 PR 30-MAY-1997; 97US-0048131.  
 PR 30-MAY-1997; 97US-0048135.  
 PR 30-MAY-1997; 97US-0048154.  
 PR 30-MAY-1997; 97US-0048160.

PR 30-MAY-1997; 97US-0048186.  
 PR 30-MAY-1997; 97US-0048187.  
 PR 30-MAY-1997; 97US-0048188.  
 PR 30-MAY-1997; 97US-0048350.  
 PR 30-MAY-1997; 97US-0048351.  
 PR 30-MAY-1997; 97US-0048352.  
 PR 30-MAY-1997; 97US-0048355.  
 PR 05-AUG-1997; 97US-0054804.

XX (HUMA-) HUMAN GENOME SCI INC.  
 XX

PI Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;  
 PI Greene JM, Hu JS, Lafleur DW, Moore PA, NI, J, Olsen HS;  
 PI Rosen CA, Ruben SM, Shi Y, Young P;  
 XX

DR WPI: 1999-070066/06.  
 DR N-PSDB: AAX00686.

XX  
 PT New isolated human genes and the secreted polypeptides they encode  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders

PS Claim 11: Page 313-315; 385pp; English.

XX  
 CC This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAX00602) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 87 novel genes and their fragments (nucleic  
 CC acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 87  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX00611 for described uses).

XX SQ Sequence 625 AA;

Query Match 6.2%; Score 38; DB 20; Length 625;

Best Local Similarity 100.0%; Pred. No. 3e-29;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 20 YQAHVSRNRKGVVTRGFRGCTVMTGLSGAGKTT 57

|||||  
 Db 30 YQAHVSRNRKGVVTRGFRGCTVMTGLSGAGKTT 67

RESULT 13

AAW75032

ID AAW75032 standard; Protein; 36 AA.

XX AC AAW75032;

XX DT 25-JAN-1999 (first entry)

XX DE Fragment of human secreted protein encoded by gene 14.

XX KW Human; secreted protein; testis; tumour; foetal brain tissue;  
 KW fusion protein; cancer; central nervous system; seizure;  
 KW diagnosis; neurodegenerative disease.

XX OS Homo sapiens.

XX PN WO9839448-A2.

XX PD 11-SEP-1998.

XX PF 06-MAR-1998; 98WO-US04493.

XX PR 02-OCT-1997; 97US-0061060.

PR 07-MAR-1997; 97US-0038621.  
 PR 07-MAR-1997; 97US-0040161.  
 PR 07-MAR-1997; 97US-0040162.  
 PR 07-MAR-1997; 97US-0040163.  
 PR 07-MAR-1997; 97US-0040333.  
 PR 07-MAR-1997; 97US-0040334.  
 PR 07-MAR-1997; 97US-0040336.  
 PR 07-MAR-1997; 97US-0040626.  
 PR 11-APR-1997; 97US-0043311.  
 PR 11-APR-1997; 97US-0043312.  
 PR 11-APR-1997; 97US-0043313.  
 PR 11-APR-1997; 97US-0043314.  
 PR 11-APR-1997; 97US-0043568.  
 PR 11-APR-1997; 97US-0043569.  
 PR 11-APR-1997; 97US-0043576.  
 PR 11-APR-1997; 97US-0043578.  
 PR 11-APR-1997; 97US-0043580.  
 PR 11-APR-1997; 97US-0043589.  
 PR 11-APR-1997; 97US-0043670.  
 PR 11-APR-1997; 97US-0043671.  
 PR 11-APR-1997; 97US-0043672.  
 PR 11-APR-1997; 97US-0043674.  
 PR 23-MAY-1997; 97US-0047492.  
 PR 23-MAY-1997; 97US-0047500.  
 PR 23-MAY-1997; 97US-0047501.  
 PR 23-MAY-1997; 97US-0047502.  
 PR 23-MAY-1997; 97US-0047503.  
 PR 23-MAY-1997; 97US-0047581.  
 PR 23-MAY-1997; 97US-0047582.  
 PR 23-MAY-1997; 97US-0047583.  
 PR 23-MAY-1997; 97US-0047584.  
 PR 23-MAY-1997; 97US-0047585.  
 PR 23-MAY-1997; 97US-0047586.  
 PR 23-MAY-1997; 97US-0047587.  
 PR 23-MAY-1997; 97US-0047588.  
 PR 23-MAY-1997; 97US-0047589.  
 PR 23-MAY-1997; 97US-0047590.  
 PR 23-MAY-1997; 97US-0047592.  
 PR 23-MAY-1997; 97US-0047593.  
 PR 23-MAY-1997; 97US-0047594.  
 PR 23-MAY-1997; 97US-0047595.  
 PR 23-MAY-1997; 97US-0047596.  
 PR 23-MAY-1997; 97US-0047597.  
 PR 23-MAY-1997; 97US-0047598.  
 PR 23-MAY-1997; 97US-0047599.  
 PR 23-MAY-1997; 97US-0047600.  
 PR 23-MAY-1997; 97US-0047601.  
 PR 23-MAY-1997; 97US-0047612.  
 PR 23-MAY-1997; 97US-0047613.  
 PR 23-MAY-1997; 97US-0047614.  
 PR 23-MAY-1997; 97US-0047615.  
 PR 23-MAY-1997; 97US-0047617.  
 PR 23-MAY-1997; 97US-0047618.  
 PR 23-MAY-1997; 97US-0047632.  
 PR 23-MAY-1997; 97US-0047633.  
 PR 06-JUN-1997; 97US-0048964.  
 PR 06-JUN-1997; 97US-0048974.  
 PR 13-JUN-1997; 97US-0049610.  
 PR 08-JUL-1997; 97US-0051926.  
 PR 16-JUL-1997; 97US-0052874.  
 PR 18-AUG-1997; 97US-0055724.  
 PR 22-AUG-1997; 97US-0056630.  
 PR 22-AUG-1997; 97US-0056631.  
 PR 22-AUG-1997; 97US-0056632.  
 PR 22-AUG-1997; 97US-0056636.  
 PR 22-AUG-1997; 97US-0056637.  
 PR 22-AUG-1997; 97US-0056662.  
 PR 22-AUG-1997; 97US-0056664.  
 PR 22-AUG-1997; 97US-0056845.  
 PR 22-AUG-1997; 97US-0056862.  
 PR 22-AUG-1997; 97US-0056864.  
 PR 22-AUG-1997; 97US-0056872.  
 PR 22-AUG-1997; 97US-0056874.

PR 22-AUG-1997; 97US-0056875.  
 PR 22-AUG-1997; 97US-0056876.  
 PR 22-AUG-1997; 97US-0056877.  
 PR 22-AUG-1997; 97US-0056878.  
 PR 22-AUG-1997; 97US-0056879.  
 PR 22-AUG-1997; 97US-0056880.  
 PR 22-AUG-1997; 97US-0056881.  
 PR 22-AUG-1997; 97US-0056882.  
 PR 22-AUG-1997; 97US-0056884.  
 PR 22-AUG-1997; 97US-0056886.  
 PR 22-AUG-1997; 97US-0056887.  
 PR 22-AUG-1997; 97US-0056888.  
 PR 22-AUG-1997; 97US-0056889.  
 PR 22-AUG-1997; 97US-0056892.  
 PR 22-AUG-1997; 97US-0056893.  
 PR 22-AUG-1997; 97US-0056894.  
 PR 22-AUG-1997; 97US-0056903.  
 PR 22-AUG-1997; 97US-0056908.  
 PR 22-AUG-1997; 97US-0056909.  
 PR 22-AUG-1997; 97US-0056910.  
 PR 22-AUG-1997; 97US-0056911.  
 PR 05-SEP-1997; 97US-0057650.  
 PR 05-SEP-1997; 97US-0057669.  
 PR 05-SEP-1997; 97US-0057761.  
 PR 12-SEP-1997; 97US-0058785.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;  
 PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;  
 PI Kyaw H, Lafleur DW, Li Y, Moore FA, Ni J, Olsen HS, Rosen CA;  
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
 XX  
 DR WPI; 1998-506364/43.  
 XX  
 XX New isolated human genes and the secreted polypeptide(s) they encode  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 PS Disclosure; Page 14; 721pp; English.  
 XX  
 XX This sequence represents a fragment of a secreted human protein encoded  
 CC by the nucleic acid molecule designated Gene 14 (AAV59524).  
 CC The gene can be used to generate fusion proteins by linking to the gene  
 CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 186 novel genes and their fragments (nucleic  
 CC acid sequences: AAV59511-V59812; amino acid sequences AAV74731-W75026)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 186  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAV59511 for described uses).  
 XX  
 SQ Sequence 36 AA;  
 Query Match 5.9%; Score 36; DB 19; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-28;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 20 YQAHVSRNKRQGVVGRGFRGCTVWLTLGLSGAGK 55  
 Db 1 YQAHVSRNKRQGVVGRGFRGCTVWLTLGLSGAGK 36  
 RESULT 14  
 AAG44041  
 ID AAG44041 standard; Protein: 373 AA.  
 XX  
 AC AAG44041;  
 XX

DT 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 55117.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS  
 XX EP1033405-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.

```
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score 22; DB 21; Length 373;
Best Local Similarity 100.0%; Pred. No. 2.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVNHGHALLM 422
Db 149 NADAVFAFQLRNPVNHGHALLM 170

RESULT 15
AAG44040
ID AAG44040 standard; Protein; 460 AA.
XX
AC AAG44040;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55116.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
```

XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX XX  
XX PD 06-SEP-2000. XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 25-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.

```
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 3.6%; Score: 22; DB 21; Length 460;

Best Local Similarity 100.0%; Fred. No. 3.4e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAQLRNPVHNGHALLM 422

|||||

Db 236 NADAVFAQLRNPVHNGHALLM 257

Search completed: May 30, 2003, 12:30:13  
Job time : 75 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:29:00 ; Search time 27 Seconds  
(without alignments)  
670.189 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 615  
Sequence: 1 MSGIKKQKTENQOKSTNVV.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 10

Total number of hits satisfying chosen parameters: 14

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	8.9	174	4	US-09-149-476-635
2	38	6.2	624	2	US-08-879-561-3
3	38	6.2	624	2	US-08-879-561-9
4	36	5.9	36	4	US-09-149-476-634
5	22	3.6	110	4	US-09-346-408-10
6	22	3.6	133	4	US-09-346-408-11
7	22	3.6	463	4	US-09-346-408-11
8	22	3.6	465	4	US-09-346-408-4
9	22	3.6	465	4	US-09-346-408-6
10	22	3.6	610	2	US-08-879-561-10
11	21	3.4	461	4	US-09-346-408-8
12	21	3.4	476	4	US-09-346-408-12
13	12	2.0	27	4	US-09-149-476-637
14	11	1.8	204	4	US-09-134-001C-4517

#### ALIGNMENTS

RESULT 1  
US-09-149-476-635  
; Sequence 635, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 8.9%; Score 55; DB 4: Length 174;

Best Local Similarity 100.0%; Pred. No. 6.1e-46; Mismatches 0; Indels 0; Gaps 0;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGIKKQKTENQKSTNVVYQAHHSRNRKRGQVVGTRGGFRGCTVWLTLGSGAGK 55  
|||||

Db 55 MSGIKKQKTENQKSTNVVYQAHHSRNRKRGQVVGTRGGFRGCTVWLTLGSGAGK 109  
|||||

# RESULT 2

US-08-879-561-3  
Sequence 3, Application US/08879561  
Patent No. 5817482  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/879,561  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0325 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LUNGN0T02  
CLONE: 373887  
US-08-879-561-3

Query Match 6.2%; Score 38; DB 2; Length 624;  
Best Local Similarity 100.0%; Pred. No. 8e-29;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 57  
Db 30 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 67

RESULT 3  
US-08-879-561-9  
Sequence 9, Application US/08879561  
Patent No. 5817482  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/879,561  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0325 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1109676  
US-08-879-561-9

Query Match 6.2%; Score 38; DB 2; Length 624;  
Best Local Similarity 100.0%; Pred. No. 8e-29;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 57  
Db 30 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 67

RESULT 4  
US-09-149-476-634  
Sequence 634, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503

[illegible]

; SEQ ID NO 2

Matches 22; Conser

Query Match 3.6%; Score 22; DB 4; Length 465;  
Best Local Similarity 100.0%; Pred. No. 2.9e-13;  
Matches 22: Conservative 0; Mismatches 0; Indels

QY 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 241 NADAVFAFQLRNPVHNGHALLM 262

RESULT 9  
US-09-346-408-6  
; Sequence 6, Application US/09346408B  
; Patent No. 6338966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01  
; EARLIER APPLICATION NUMBER: 60/092,833  
; EARLIER FILING DATE: July 14, 1998  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 465  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-346-408-6

Query Match 3.6%; Score 22; DB 4; Length 465;  
Best Local Similarity 100.0%; Pred. No. 2.9e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 241 NADAVFAFQLRNPVHNGHALLM 262

RESULT 10  
US-08-879-561-10  
; Sequence 10, Application US/08879561  
; Patent No. 5817482  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Hawkins, Phillip R.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/879,561  
; FILING DATE: Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0325 US  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 705385  
US-08-879-561-10

Query Match 3.6%; Score 22; DB 2; Length 610;  
Best Local Similarity 100.0%; Pred. No. 3.6e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 396 NADAVFAFQLRNPVHNGHALLM 417

RESULT 11  
US-09-346-408-8  
; Sequence 8, Application US/09346408B  
; Patent No. 6338966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01  
; EARLIER APPLICATION NUMBER: 60/092,833  
; EARLIER FILING DATE: July 14, 1998  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 8  
; LENGTH: 461  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
US-09-346-408-8

Query Match 3.4%; Score 21; DB 4; Length 461;  
Best Local Similarity 100.0%; Pred. No. 2.7e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 ADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 239 ADAVFAFQLRNPVHNGHALLM 259

RESULT 12  
US-09-346-408-12  
; Sequence 12, Application US/09346408B  
; Patent No. 6338966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01  
; EARLIER APPLICATION NUMBER: 60/092,833  
; EARLIER FILING DATE: July 14, 1998  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 12

LENGTH: 476  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-346-408-12

Query Match 3.4%; Score 21; DB 4; Length 476;  
Best Local Similarity 100.0%; Pred. No. 2.8e-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 ADAVFAFOLRNPVNGHALLM 422  
|||||  
DB 254 ADAVFAFOLRNPVNGHALLM 274

RESULT 13

US-09-149-476-637  
Sequence 637, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: PZ002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,910  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,864  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,631  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,845  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,892  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,761  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/047,595  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,599  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,588  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,585  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,586  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,590  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,594  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,589  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,593  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,614  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,578  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,576  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/047,501  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,670  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/056,632  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,664  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,876  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,881  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,909  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,875  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,862  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,887  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,908  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/048,964  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/057,650  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/056,884  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,669  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/049,610  
 ; EARLIER FILING DATE: 1997-06-13  
 ; EARLIER APPLICATION NUMBER: 60/061,060  
 ; EARLIER FILING DATE: 1997-10-02

Query Match 2.0%; Score 12; DB 4; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 76 LDGDNVRRHGLNR 87  
 |||||  
 Db 2 LDGDNVRRHGLNR 13  
 |||||  
 RESULT 14  
 US-09-134-001C-4517  
 ; Sequence 4517, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 4517  
 ; LENGTH: 204  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-4517

Query Match 1.8%; Score 11; DB 4; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 0.0083;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRRHGLN 86  
 |||||  
 Db 66 LDGDNVRRHGLN 76

Search completed: May 30, 2003, 12:33:43  
 Job time : 28 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:32:25 ; Search time 49 Seconds  
(without alignments)  
1270.461 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 615

Sequence: 1 MSGIKKOKTENQOKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 393519 seqs, 10123694 residues

Word size : 10

Total number of hits satisfying chosen parameters: 31

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	221	35.9	619	9	US-09-974-298-184
2	65	10.6	65	10	US-09-864-761-39981
3	55	8.9	174	9	US-09-809-391-635
4	38	6.2	624	9	US-09-984-245-196
5	38	6.2	624	9	US-09-966-262-196
6	38	6.2	624	9	US-09-983-966-196
7	38	6.2	624	9	US-10-143-090-196
8	38	6.2	624	10	US-09-771-161A-244
9	36	5.9	36	9	US-09-809-391-634
10	28	4.6	45	9	US-09-984-245-332
11	28	4.6	45	9	US-09-966-262-332
12	28	4.6	45	9	US-09-983-966-332
13	28	4.6	45	9	US-10-143-090-332
14	20	3.3	41	9	US-09-984-245-340
15	20	3.3	41	9	US-09-966-262-340
16	20	3.3	41	9	US-09-983-966-340
17	20	3.3	41	9	US-10-143-090-340
18	17	2.8	37	9	US-09-984-245-339
19	17	2.8	37	9	US-09-966-262-339

RESULT 1  
US-09-974-298-184  
; Sequence 184, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hwei-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 184  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1420940CD1  
US-09-974-298-184

Query Match 35.9%; Score 221; DB 9; Length 619;  
Best Local Similarity 99.7%; Pred. No. 1.6e-206;  
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	289	DGVNMSIPVLPVSAEDKTRLEGSKFVLAHGGRRVAILRDAEFYHRKEERCRCRVWGT	348
Db	294	DGVNMSIPVLPVSAEDKTRLEGSKFVLAHGGRRVAILRDAEFYHRKEERCRCRVWGT	353
Qy	349	TCTKHPHKVMWESGDLVGGDLQVLEKIRWNDGLDQVRLTPLELKQCKEMNADAVAF	408
Db	354	TCTKHPHKVMWESGDLVGGDLQVLEKIRWNDGLDQVRLTPLELKQCKEMNADAVAF	413
Qy	409	QLRNPVNHGALLMQDTCRRLLRGYKHPVLLLLHPLGWTDDDDVPLDWRMKQHAALVEE	468
Db	414	QLRNPVNHGALLMQDTCRRLLRGYKHPVLLLLHPLGWTDDDDVPLDWRMKQHAALVEE	473
Qy	469	GVLPDKSTIVAIFPSPMLYAGTPEVQWCHRCRMTAGANFYIVGRDPAGMPHPETKKDLYE	528
Db	474	GVLPDKSTIVAIFPSPMLYAGTPEVQWCHRCRMTAGANFYIVGRDPAGMPHPETKKDLYE	533
Qy	529	PTHGKVLSMAPGLTSVEIIPFVAAYNKAAMDYDPARHNEFDFTSGTRMRKLAREG	588
Db	534	PTHGKVLSMAPGLTSVEIIPFVAAYNKAAMDYDPARHNEFDFTSGTRMRKLAREG	593
Qy	589	ENPPDGFMAPKAWKVLTDYRS 610	
Db	594	ENPPDGFMAPKAWKVLTDYRS 615	

RESULT 2  
US-09-864-761-39981  
; Sequence 39981, Application US/09864761  
; Patent No. US20020048763A1

GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39981  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Homo sapiens  
OTHER INFORMATION: MAP TO AC006191.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
OTHER INFORMATION: SWISSPROT HIT: O95340, EVALUATE 4.00e-33  
OTHER INFORMATION: EST\_HUMAN HIT: BE292722.1, EVALUATE 4.00e-32

US-09-864-761-39981  
Query Match 10.6%; Score 65; DB 10; Length 65;  
Best Local Similarity 100.0%; Pred. No. 1.6e-55;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 63 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDREINIRIAEVAKLFADAGLVCTTST 122  
DB 1 EYLVSHAIPCVSLDGNVHGLNRLNLFSPGDREINIRIAEVAKLFADAGLVCTTST 60

QY 123 SPFAK 127  
DB 61 SPFAK 65  
RESULT 3  
US-09-809-391-635  
Sequence 635, Application US/09809391  
Publication No. US20030049618A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 635  
LENGTH: 174  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-809-391-635  
Query Match 8.9%; Score 55; DB 9; Length 174;  
Best Local Similarity 100.0%; Pred. No. 2e-45; Indels 0; Gaps 0;  
Matches 55; Conservative 0; Mismatches 0;  
QY 1 MSGIKKQKTENQOKSTNVVYQAHVSRNKRQGVVGTGFGTGVLTGLSGAGK 55  
DB 55 MSGIKKQKTENQOKSTNVVYQAHVSRNKRQGVVGTGFGTGVLTGLSGAGK 109

RESULT 4  
US-09-984-245-196  
Sequence 196, Application US/09984245  
Patent No. US20020165374A1  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 87 Human Secreted Proteins  
FILE REFERENCE: P2004P1  
CURRENT APPLICATION NUMBER: US/09/984,245  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 09/154,707  
PRIOR FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: PCT/US98/05311  
PRIOR FILING DATE: 1998-03-19  
PRIOR APPLICATION NUMBER: US 60/041,277  
PRIOR FILING DATE: 1997-03-21  
PRIOR APPLICATION NUMBER: US 60/042,344  
PRIOR FILING DATE: 1997-03-21  
PRIOR APPLICATION NUMBER: US 60/041,276  
PRIOR FILING DATE: 1997-03-21  
PRIOR APPLICATION NUMBER: US 60/041,281  
PRIOR FILING DATE: 1997-03-21  
PRIOR APPLICATION NUMBER: US 60/048,094  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,350  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,188  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,135  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/050,937  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,187  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,099  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,352  
PRIOR FILING DATE: 1997-05-30  
PRIOR APPLICATION NUMBER: US 60/048,186  
PRIOR FILING DATE: 1997-05-30

RESULT 5  
US-09-366-262-196  
; Sequence 196, Application US/09966262  
; Publication No. US20030050461A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Pro  
; FILE REFERENCE: P2004P1  
; CURRENT APPLICATION NUMBER: US/09/966,262  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: US 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30

```

RESULT 6
US-09-983-966-196
; Sequence 196, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30

```

;; PRIOR APPLICATION NUMBER: US 60/048,135  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/050,937  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,187  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,099  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,352  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,186  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,069  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,095  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,131  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,096  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,355  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/054,804  
;; PRIOR FILING DATE: 1997-08-05  
;; PRIOR APPLICATION NUMBER: US 60/056,370  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: US 60/060,862  
;; PRIOR FILING DATE: 1997-10-02  
;; NUMBER OF SEQ ID NOS: 343  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 196  
;; LENGTH: 624  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-983-966-196

Query Match 6.2%; Score 38; DB 9; Length 624;  
Best Local Similarity 100.0%; Pred. No. 2.1e-28;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 57  
|||||  
DB 30 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 67.

RESULT 7  
US-10-143-090-196  
;; Sequence 196, Application US/10143090  
;; Publication No. US20030069406A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Young et al.  
;; TITLE OF INVENTION: 87 Human Secreted Proteins  
;; FILE REFERENCE: P2004P1  
;; CURRENT APPLICATION NUMBER: US/10/143,090  
;; CURRENT FILING DATE: 2002-05-13  
;; PRIOR APPLICATION NUMBER: 09/154,707  
;; PRIOR FILING DATE: 1998-09-17  
;; NUMBER OF SEQ ID NOS: 343  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 196  
;; LENGTH: 624  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-143-090-196

Query Match 6.2%; Score 38; DB 9; Length 624;  
Best Local Similarity 100.0%; Pred. No. 2.1e-28;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 20 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 57  
|||||  
DB 30 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 67  
RESULT 8  
US-09-771-161A-244  
;; Sequence 244, Application US/09771161A  
;; Patent No. US20020110811A1  
;; GENERAL INFORMATION:  
;; APPLICANT: LEVINE, et al.  
;; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
;; FILE REFERENCE: 802620-2005.1  
;; CURRENT APPLICATION NUMBER: US/09/771.161A  
;; CURRENT FILING DATE: 2001-01-26  
;; PRIOR APPLICATION NUMBER: 09/724,676  
;; PRIOR FILING DATE: 2000-11-28  
;; PRIOR APPLICATION NUMBER: 136776  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: 135619  
;; PRIOR FILING DATE: 2000-04-12  
;; NUMBER OF SEQ ID NOS: 273  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 244  
;; LENGTH: 624  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-771-161A-244

Query Match 6.2%; Score 38; DB 10; Length 624;  
Best Local Similarity 100.0%; Pred. No. 2.1e-28;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 57  
|||||  
DB 30 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGKTT 67

RESULT 9  
US-09-809-391-634  
;; Sequence 634, Application US/09809391  
;; Publication No. US20030049618A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 186 Human Secreted proteins  
;; FILE REFERENCE: P2002P2  
;; CURRENT APPLICATION NUMBER: US/09/809,391  
;; CURRENT FILING DATE: 2001-03-16  
;; Prior application data removed - consult PALM or file wrapper  
;; NUMBER OF SEQ ID NOS: 761  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 634  
;; LENGTH: 36  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-809-391-634

Query Match 5.9%; Score 36; DB 9; Length 36;  
Best Local Similarity 100.0%; Pred. No. 1.6e-27;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGK 55  
|||||  
DB 1 YQAHVSRNKRGOVGTGRGCTVWLTGLSGAGK 36

RESULT 10  
US-09-984-245-332  
;; Sequence 332, Application US/09984245  
;; Patent No. US20020165374A1  
;; GENERAL INFORMATION:

```

; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-332

```

Query Match 4.6%; Score 28; DB 9; Length 45;

Best Local Similarity 100.0%; Pred. No. 1.2e-19; Indels 0; Gaps 0;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 YQAHVSRNKRQGVVGTGRGCTVWL 47  
 |||||  
 Db 18 YQAHVSRNKRQGVVGTGRGCTVWL 45

```

RESULT 11
US-09-966-262-332
; Sequence 332, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-262-332

```

Query Match 4.6%; Score 28; DB 9; Length 45;

Best Local Similarity 100.0%; Pred. No. 1.2e-19;

Matches	28;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

**Qy** 20 YQAHHVSRNKRQQVVGTRGGFRGCTVWL 47  
+ + + + + + + + + + + + + + + + + +  
**Db** 18 YQAHHVSRNKRQQVVGTRGGFRGCTVWL 45

## RESULT. 12

```

US-09-983-966-332
; Sequence 332, Application US/09983966
; Publication No. US2003006019A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004F1
; CURRENT APPLICATION NUMBER: US/09/983,966
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 332
; LENGTH: 45

```

RESULT 13

```

US-10-143-090-332
; Sequence 332, Application US/10143090
; Publication No. US20030069406A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/10/143,090
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-090-332

```

## RESULT 14

```

US-09/984-245-340
; Sequence 340, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: PZ004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937

```

; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,154  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: US 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: US 60/060,862  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 340  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-245-340

Query Match 3.3%; Score 20; DB 9; Length 41;  
Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AGANFYIVGRDPAGMPHPET 522  
Db 18 AGANFYIVGRDPAGMPHPET 37  
|||||

## RESULT 15

US-09-966-262-340  
; Sequence 340, Application US/09966262  
; Publication No. US20030050461A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004p1  
; CURRENT APPLICATION NUMBER: US/09/966,262  
; PRIOR FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: US 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350

; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,154  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: US 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: US 60/060,862  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 340  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-966-262-340

Query Match 3.3%; Score 20; DB 9; Length 41;  
Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AGANFYIVGRDPAGMPHPET 522  
Db 18 AGANFYIVGRDPAGMPHPET 37  
|||||

Search completed: May 30, 2003, 12:41:34  
Job time : 50 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:28:05 ; Search time 43 Seconds  
(without alignments)  
1374.946 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 615  
Sequence: 1 MSGIKKKTENQOKSTNVY.....MAPKAWLVLDYRSEMDKN 615

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 283224 seqs, 96134422 residues  
Word size : 10

Total number of hits satisfying chosen parameters: 32

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR-73.\*

- 1: pir1.\*
- 2: pir2.\*
- 3: pir3.\*
- 4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 38    | 6.2         | 624    | 1 JW0087 | adenylyl-sulfate k |
| 2          | 22    | 3.6         | 424    | 2 S44079 | sulfate adenylyltr |
| 3          | 22    | 3.6         | 463    | 2 S44267 | sulfate adenylyltr |
| 4          | 22    | 3.6         | 489    | 1 T01204 | sulfate adenylyltr |
| 5          | 22    | 3.6         | 610    | 1 J04383 | adenylyl-sulfate k |
| 6          | 21    | 3.4         | 459    | 2 T08594 | probable sulfate a |
| 7          | 21    | 3.4         | 465    | 2 E71409 | sulfate adenylyltr |
| 8          | 21    | 3.4         | 469    | 2 T52659 | sulfate adenylyltr |
| 9          | 21    | 3.4         | 476    | 2 S44943 | sulfate adenylyltr |
| 10         | 21    | 3.4         | 483    | 2 T14475 | sulfate adenylyltr |
| 11         | 18    | 2.9         | 652    | 1 T24918 | 3'-phosphoadenosin |
| 12         | 13    | 2.1         | 177    | 1 S74917 | adenylyl-sulfate k |
| 13         | 13    | 2.1         | 202    | 2 T50101 | adenylylsulfate ki |
| 14         | 12    | 2.0         | 276    | 1 S47640 | adenylyl-sulfate k |
| 15         | 12    | 2.0         | 293    | 2 T06100 | adenylyl-sulfate k |
| 16         | 12    | 2.0         | 633    | 2 H95932 | probable adenylyl- |
| 17         | 12    | 2.0         | 641    | 1 Z22RNQ | adenylyl-sulfate k |
| 18         | 12    | 2.0         | 641    | 2 E95320 | adenylyl-sulfate k |
| 19         | 12    | 2.0         | 660    | 2 G82672 | ATP sulfurylase, l |
| 20         | 11    | 1.8         | 110    | 2 C55228 | hypothetical prote |
| 21         | 11    | 1.8         | 196    | 2 H83472 | adenosine 5'-phosp |
| 22         | 11    | 1.8         | 214    | 1 S18729 | adenylyl-sulfate k |
| 23         | 11    | 1.8         | 312    | 2 T08076 | adenylyl-sulfate k |
| 24         | 11    | 1.8         | 635    | 2 A87433 | hypothetical prote |
| 25         | 11    | 1.8         | 644    | 2 AD3471 | adenylyl-sulfate k |
| 26         | 10    | 1.6         | 200    | 2 E96912 | adenylylsulfate ki |
| 27         | 10    | 1.6         | 201    | 1 B65056 | adenylyl-sulfate k |
| 28         | 10    | 1.6         | 201    | 2 D91079 | adenosine 5'-phosp |
| 29         | 10    | 1.6         | 201    | 2 E85924 | adenosine 5'-phosp |

30 10 1.6 201 2 AH0856 adenosine 5-phosph  
31 10 1.6 213 2 AF0408 adenylyl-sulfate k  
32 10 1.6 215 2 F82062 adenylylsulfate ki

ALIGNMENTS

RESULT 1  
JW0087

adenylyl-sulfate kinase (EC 2.7.1.25) - human  
N:Alternate names: adenosine 5'-phosphosulfate kinase: PAPS  
N:Contains: adenylylsulfate kinase (EC 2.7.1.25); sulfate adenylyltransferase (EC 2.7.1.25)  
C:Species: Homo sapiens (man)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002  
C:Accession: JW0087  
R:Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakayama, T.; Nakajima, H.; Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998  
A:Title: cDNA cloning, expression, and characterization of the human bifunctional A.P  
A:Reference number: JW0087; MUID:98312048; PMID:9648242  
A:Accession: JW0087  
A:Molecule type: mRNA  
A:Residues: 1-624 <YAN>  
A:Cross-references: GB:AF033026; NID:g3378100; PIDN:AAC28429.1; PID:g3378101  
A:Experimental source: Brain  
C:Function: <ASKF>  
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsul  
C:Function: <SATF>  
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and A  
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfa  
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loo  
F:52-215/Domain: adenylylsulfate kinase homology <ASK>  
F:59-66/Region: nucleotide-binding motif A (P-loop)  
F:226-620/Domain: sulfate adenylyltransferase homology <SAT>  
F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 6.2%; Score 38; DB 1; Length 624;  
Best Local Similarity 100.0%; Pred. No. 2e-30; 0; Indels 0; Gaps 0;  
Matches 38; Conservative 0; Mismatches 0;

QY 20 YQAHVSRNKRGOVVGTRGGRGCTVWLTGLSGAGKTT 57  
|||||  
DB 30 YQAHVSRNKRGOVVGTRGGRGCTVWLTGLSGAGKTT 67  
|||||

RESULT 2  
S44079  
sulfate adenylyltransferase (EC 2.7.7.4) met3-1 - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: S67497; S44079  
R:Klonus, D.; Hoefgen, R.; Willmitzer, L.; Riesmeier, J.W.  
Plant J. 6, 105-112, 1994  
A:Title: Isolation and characterization of two cDNA clones encoding ATP-sulfurylas  
A:Reference number: S67497; MUID:95004649; PMID:7920699  
A:Accession: S67497  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-424 <KLO>  
A:Cross-references: EMBL:X75041; NID:g471342; PIDN:CAA52953.1; PID:g531495  
C:Genetics: met3-1  
A:Gene: met3-1

C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homo-  
C:Keywords: nucleotidyltransferase  
F:11-409/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.6%; Score 22; DB 2; Length 424;  
Best Local Similarity 100.0%; Pred. No. 4.6e-14; 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0;

QY 401 NADAVFAFQLRNPVINGHALLM 422  
|||||  
DB 200 NADAVFAFQLRNPVINGHALLM 221  
|||||

## RESULT 3

S44267  
sulfate adenylyltransferase (EC 2.7.7.4) met3-2 - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 08-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: S67498; S44267  
R:Klonus, D.; Hoefgen, R.; Willmitzer, L.; Riesmeier, J.W.  
Plant J. 6, 105-112, 1994  
A:Title: Isolation and characterization of two cDNA clones encoding ATP-sulfurylases from potato  
A:Reference number: S67497; MUID:95004649; PMID:7920699  
A:Accession: S67498  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-463 <KLO>  
A:Cross-references: EMBL:X79053; NID:g479089; PIDN:CAA55655.1; PID:g479090  
C:Genetics:  
A:Gene: met3-2  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: chloroplast; nucleotidyltransferase  
F:50-448/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.6%; Score 22; DB 2; Length 463;  
Best Local Similarity 100.0%; Pred. No. 5e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 239 NADAVFAFQLRNPVHNGHALLM 260

## RESULT 4

T01204  
sulfate adenylyltransferase (EC 2.7.7.4) - maize  
N:Alternate names: ATP sulfurylase  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 23-Mar-2001  
C:Accession: T01204  
R:Boichi, A.; Petrucco, S.; Foroni, C.; Tencza, G.L.; Ottonello, S.  
submitted to the EMBL Data Library, July 1997  
A:Description: Sulfate permease and ATP sulfurylase mRNA are coordinately modulated in maize  
A:Reference number: Z14260  
A:Accession: T01204  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-489 <BOL>  
A:Cross-references: EMBL:AF016305; NID:g2738749; PIDN:AAB94542.1; PID:g2738750  
A:Experimental source: cultivar Paulo; root  
C:Function:

A:Description: catalyzes activation of sulfate to adenylylsulfate  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: nucleotidyltransferase  
F:70-473/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.6%; Score 22; DB 1; Length 489;  
Best Local Similarity 100.0%; Pred. No. 5.3e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 264 NADAVFAFQLRNPVHNGHALLM 285

## RESULT 5

JC4383  
adenylyl-sulfate kinase (EC 2.7.1.25) - spoonworm (Urechis caupo)  
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS  
N:Contains: adenylyl-sulfate kinase (EC 2.7.1.25); sulfate adenylyltransferase (EC 2.7.7.4)  
C:Species: Urechis caupo  
C:Date: 17-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 03-Jun-2002  
C:Accession: JC4383  
R:Rosenthal, E.; Leustek, T.

Gene 165, 243-248, 1995  
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfu  
A:Reference number: JC4383; MUID:96096529; PMID:8522184  
A:Accession: JC4383

A:Molecule type: mRNA  
A:Residues: 1-610 <ROS>  
A:Cross-references: GB:L39001; NID:g705384; PIDN:AAB00139.1; PID:g705385  
C:Genetics:  
A:Gene: papss  
C:Function: <ASKF>  
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsul  
C:Function: <SATF>  
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and A  
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfa  
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loo  
F:37-200/Domain: adenylylsulfate kinase homology <ASK>  
F:44-51/Region: nucleotide-binding motif A (P-loop)  
F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PA  
F:211-605/Domain: sulfate adenylyltransferase homology <SAT>  
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 3.6%; Score 22; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 6.5e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 396 NADAVFAFQLRNPVHNGHALLM 417

## RESULT 6

T08594  
probable sulfate adenylyltransferase (EC 2.7.7.4) - rape  
C:Species: Brassica napus (rape)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T08594  
R:Buchanan-Wollaston, V.; Ainsworth, C.  
Plant Mol. Biol. 33, 821-834, 1997  
A:Title: Leaf senescence in Brassica napus: cloning of senescence related genes by su  
A:Reference number: Z16446; MUID:97260386; PMID:9106506  
A:Accession: T08594  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-459 <BUC>  
A:Cross-references: EMBL:U68218; NID:g1527218; PIDN:AAB53100.1; PID:g1527219  
A:Experimental source: cv. Falcon; senescing leaves  
C:Genetics:

A:Gene: LSC680  
C:Function:  
A:Description: catalyzes the activation of sulfate to adenylylsulfate  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo  
C:Keywords: nucleotidyltransferase  
F:45-444/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.4%; Score 21; DB 2; Length 459;  
Best Local Similarity 100.0%; Pred. No. 5.4e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFAFQLRNPVHNGHALLM 422  
|||||  
Db 236 ADAVFAFQLRNPVHNGHALLM 256

## RESULT 7

E71409  
sulfate adenylyltransferase (EC 2.7.7.4) precursor (clone APS3) - Arabidopsis thalian  
N:Alternate names: ATP sulfurylase  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: E71409; S68201  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D  
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: E71409  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-465 <BEV>  
A:Cross-references: GB:297336; NID:92244788; PIDN:CAB10247.1; PID:92244824  
R:Murillo, M.; Leustek, T.  
Arch. Biochem. Biophys. 323, 195-204, 1995  
A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherichia  
ine-5'-triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis of a recom  
A:Reference number: S68024; MUID:96019964; PMID:7487067  
A:Accession: S68201  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-6, 'D', 8-35, 'A', 37-465 <MUR>  
A:Cross-references: EMBL:U06275; NID:9459143; PIDN:AAA923350.1; PID:9459144  
C:Genetics:  
A:Map position: 4COP9-4G3845  
A:Genome: nuclear  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: chloroplast; nucleotidyltransferase  
F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:51-465/Product: sulfate adenylyltransferase #status predicted <MAT>  
F:52-450/Domain: sulfate adenylyltransferase homology <SAT>  
  
Query Match 3.4%; Score 21; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 5.4e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 402 ADAVFATQLRNPVHNGHALM 422  
Db 242 ADAVFATQLRNPVHNGHALM 262  
|||||  
  
RESULT 8  
T52659  
sulfate adenylyltransferase (EC 2.7.7.4) aps4 precursor [validated] - Arabidopsis thalia  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 03-Nov-2000  
C:Accession: T52659  
R:Hatzfeld, Y.; Lee, S.; Lee, M.; Leustek, T.; Saito, K.  
Gene 248, 51-58, 2000  
A:Title: Functional characterization of a gene encoding a fourth ATP sulfurylase isoform  
A:Reference number: 226165  
A:Accession: T52659  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-469 <HAT>  
A:Cross-references: EMBL:AJ012586; PIDN:CAB42640.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: aps4  
A:Genome: nuclear  
C:Function:  
A:Description: EC 2.7.7.4 [validated, MUID:20267847]  
A:Pathway: sulfate assimilation  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: chloroplast; nucleotidyltransferase  
  
Query Match 3.4%; Score 21; DB 2; Length 469;  
Best Local Similarity 100.0%; Pred. No. 5.5e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 402 ADAVFATQLRNPVHNGHALM 422  
Db 244 ADAVFATQLRNPVHNGHALM 264  
|||||

## RESULT 9

S44943  
sulfate adenylyltransferase (EC 2.7.7.4) met3-1 precursor - Arabidopsis thaliana  
N:Alternate names: ATP sulfurylase; sulfate adenylyltransferase APS2  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 06-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: S44943; S68024; S68202  
R:Klonus, D.  
submitted to the EMBL Data Library, May 1994  
A:Description: ATP-sulfurylase cDNA clone from A.thaliana.  
A:Reference number: S44943  
A:Accession: S44943  
A:Molecule type: mRNA  
A:Residues: 1-476 <KLO>  
A:Cross-references: EMBL:X79210; NID:9487403; PIDN:CAA55799.1; PID:9487404  
R:Murillo, M.; Leustek, T.  
Arch. Biochem. Biophys. 323, 195-204, 1995  
A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherich  
ine-5'-triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis of a rec  
A:Reference number: S68024; MUID:96019964; PMID:7487067  
A:Accession: S68024  
A:Molecule type: mRNA  
A:Residues: 1-476 <MUW>  
A:Cross-references: EMBL:U06276; NID:9228103; PIDN:AAA92351.1; PID:9228104  
A:Experimental source: Clone APS2  
C:Genetics:  
A:Gene: met3-1  
A:Genome: nuclear  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo  
C:Keywords: chloroplast; nucleotidyltransferase  
F:1-62/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:63-476/Product: sulfate adenylyltransferase met3-1 #status predicted <MAT>  
F:64-463/Domain: sulfate adenylyltransferase homology <SAT>  
  
Query Match 3.4%; Score 21; DB 2; Length 476;  
Best Local Similarity 100.0%; Pred. No. 5.6e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 402 ADAVFATQLRNPVHNGHALM 422  
Db 254 ADAVFATQLRNPVHNGHALM 274  
|||||  
  
RESULT 10  
T14475  
sulfate adenylyltransferase (EC 2.7.7.4) ASBo precursor - wild cabbage  
N:Alternate names: ATP sulfurylase  
C:Species: Brassica oleracea (wild cabbage)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000  
C:Accession: T14475  
R:Hatzfeld, Y.; Logan, H.M.; Cathala, N.; Davidian, J.C.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z18111  
A:Accession: T14475  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-483 <HAT>  
A:Cross-references: EMBL:U06964; NID:91589912; PID:91589913  
C:Genetics:  
A:Gene: ASBo  
A:Introns: 185/2; 407/3; 436/3  
C:Function:  
A:Description: catalyzes the activation of sulfate to adenylylsulfate  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo  
C:Keywords: nucleotidyltransferase  
F:68-466/Domain: sulfate adenylyltransferase homology <STR>  
  
Query Match 3.4%; Score 21; DB 2; Length 483;  
Best Local Similarity 100.0%; Pred. No. 5.6e-13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 402 ADAVFATQLRNPVHNGHALM 422

```
|||||
258 ADAVFAPQLRPNVHNGHALLM 278

Db
RESULT 11
T24918
3'-phosphoadenosine-5'-phosphosulfate synthetase - Caenorhabditis elegans
N:Alternate names: protein T14G10.1
C:Species: Caenorhabditis elegans
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 19-Jan-2001
C:Accession: T24918
R:Wild, A.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19954
A:Accession: T24918
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-652 <WTL>
A:Cross-references: EMBL:Z68880; PIDN:CAA93098.1; GSPDB:GN00022; CESP:T14G10.1
A:Experimental source: clone T14G10.
C:Genetics:
A:Gene: CESP:T14G10.1
A:Map position: 4
A:Introns: 23/3; 82/3; 176/1; 535/2; 623/2
C:Function: <ASK>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsulfate
C:Function: <SAFE>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and ATP
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:56-219/Domain: adenylylsulfate kinase homology <ASK>
F:63-70/Region: nucleotide-binding motif A (P-loop)
F:180-190/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PAB>
F:230-647/Domain: sulfate adenylyltransferase homology <SAT>
F:137/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.9%; Score 18; DB 1; Length 652;
Best Local Similarity 100.0%; Pred. No. 9.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 LLLHPLGGWTKDDVPLD 456
|||||
Db 476 LLLHPLGGWTKDDVPLD 493

RESULT 12
S74917
adenylyl-sulfate kinase (EC 2.7.1.25) - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein slr0676
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S74917
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74917
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-177 <KAN>
A:Cross-references: EMBL:D90902; GB:AB001339; NID:g1652027; PIDN:BAAL16957.1; PID:g165203
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: cysC
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:5-164/Domain: adenylylsulfate kinase homology <ASK>
F:12-19/Region: nucleotide-binding motif A (P-loop)

Query Match 2.1%; Score 13; DB 1; Length 177;
```

```
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 WLTGLSGAGKTTI 58
|||||
Db 9 WLTGLSGAGKTTI 21

RESULT 13
T50101
adenylylsulfate kinase [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: T50101
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25037
A:Accession: T50101
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-202 <WOO>
A:Cross-references: EMBL:AL158056; PIDN:CAB76273.1; GSPDB:GN00066; SPDB:SPAC1782.11
A:Experimental source: strain 972h(-); cosmid cl782
C:Genetics:
A:Gene: SPDB:SPAC1782.11
A:Map position: 1
A:Introns: 157/1
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 2.1%; Score 13; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KGLYKRARAGEIK 173
|||||
Db 144 KGLYKRARAGEIK 156

RESULT 14
S47640
adenylyl-sulfate kinase (EC 2.7.1.25) precursor - Arabidopsis thaliana
N:Alternate names: APS kinase; protein F26C24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S47640; T02601; A84521; S38587
R:Arz, H.E.; Gisseimann, G.; Schifmann, S.; Schwenn, J.D.
Biochim. Biophys. Acta 1218, 447-452, 1994
A:Title: A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana.
A:Reference number: S47640; MUID:94325358; PMID:8049272
A:Accession: S47640
A:Molecule type: mRNA
A:Residues: 1-276 <ARZ>
A:Cross-references: EMBL:X75782; NID:g414736; PIDN:CAA53426.1; PID:g414737
R:Notes: it is uncertain whether Met-1, Met-14 or Met-26 is the initiator
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A:Reference number: Z14680
A:Accession: T02601
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-276 <ROU>
A:Cross-references: EMBL:AC004705; NID:g3252804; PIDN:AAAC24182.1; PID:g3252812
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: AB4521
A>Status: preliminary
A:Molecule type: DNA
```

A:Residues: 1-276 <STO>  
A:Cross-references: GB:AB002093; NID:g3252812; PIDN:AAC24182.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g14750; F26C24.11  
A:Map position: 2  
A:Genome: nuclear

A:Introns: 66/1; 113/1; 163/1; 202/3; 231/1; 246/3  
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology  
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding  
F:1-37/Domain: transit peptide (chloroplast) #status predicted <TNP>  
F:38-276/Product: adenylylsulfate kinase #status predicted <KAT>  
F:101-262/Domain: adenylylsulfate kinase homology <ASK>  
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 2.0%; Score 12; DB 1; Length 276;

Best Local Similarity 100.0%; Pred. No. 0.00065;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87

|||||

Db 135 LDGDNVRHGLNR 146

#### RESULT 15

T06100

adenylyl-sulfate kinase (EC 2.7.1.25) [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 03-Jun-2002

C:Accession: T06100; T52055

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15184

A:Accession: T06100

A:Molecule type: DNA

A:Residues: 1-293 <BEV>

A:Cross-references: EMBL:AL035708

A:Experimental source: cultivar Columbia; BAC clone T5J17

R:Arz, H.E.; Gisselmann, G.; Schiffmann, S.; Schwenk, J.D.

Biochim. Biophys. Acta 1218, 447-452, 1994

A:Title: A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana.

A:Reference number: S47640; MUID:94325358; PMID:8049272

A:Accession: T52055

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-293 <ARZ>

A:Cross-references: EMBL:AF043351; PIDN:AAC39520.1

C:Genetics:

A:Gene: akn2

A:Map position: 4

A:Introns: 78/1; 177/1; 216/3; 245/1; 260/3

A:Note: T5J17.110

C:Function:

A:Description: EC 2.7.1.25 [validated, MUID:94325358]

C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding

F:115-279/Domain: adenylylsulfate kinase homology <ASK>

Query Match 2.0%; Score 12; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.00069;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87

|||||

Db 149 LDGDNVRHGLNR 160

Search completed: May 30, 2003, 12:33:08

Job time : 44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:21:10 ; Search time 24 Seconds  
(without alignments)  
1062.831 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 615  
Sequence: 1 MSGIKKQKTEKQKSTNVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 10

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 509   | 82.8        | 614    | 1     | PPS2_HUMAN  |
| 2          | 81    | 13.2        | 621    | 1     | PPS2_MOUSE  |
| 3          | 38    | 6.2         | 624    | 1     | PPS1_CAVPO  |
| 4          | 38    | 6.2         | 624    | 1     | PPS1_HUMAN  |
| 5          | 38    | 6.2         | 624    | 1     | PPS1_MOUSE  |
| 6          | 15    | 2.4         | 206    | 1     | KAPS_EMENI  |
| 7          | 15    | 2.4         | 211    | 1     | KAPS_PENCH  |
| 8          | 13    | 2.1         | 177    | 1     | CYSC_SYNY3  |
| 9          | 12    | 2.0         | 276    | 1     | KAP1_ARATH  |
| 10         | 12    | 2.0         | 293    | 1     | KAP2_ARATH  |
| 11         | 12    | 2.0         | 623    | 1     | CYSN_XYLFA  |
| 12         | 12    | 2.0         | 632    | 1     | NODQ_RHITR  |
| 13         | 12    | 2.0         | 633    | 1     | NODQ_RHISB  |
| 14         | 12    | 2.0         | 641    | 1     | NODQ_RHIME  |
| 15         | 12    | 2.0         | 646    | 1     | NODQ_RHIS3  |
| 16         | 11    | 1.8         | 110    | 1     | YFMT_THETH  |
| 17         | 11    | 1.8         | 196    | 1     | CYCL_PSEAE  |
| 18         | 11    | 1.8         | 214    | 1     | CYC2_PSEAE  |
| 19         | 11    | 1.8         | 312    | 1     | KAPS_CATRO  |
| 20         | 10    | 1.6         | 200    | 1     | CYSC_ECOLI  |
| 21         | 10    | 1.6         | 215    | 1     | CYSC_VIBCH  |

#### ALIGNMENTS

RESULT 1  
PPS2\_HUMAN  
ID PPS2\_HUMAN STANDARD; 614 AA.  
AC Q95340; Q9UP30; Q9BZL2; Q9P0G6; Q9UHM1; Q9UKD3;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS  
DE synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK 2)  
DE [Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl  
DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase  
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)  
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-  
DE phosphoadenosine-5'-phosphosulfate synthetase)].  
GN PAPS2 OR ATPSK2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RC TISSUE=Fetal cartilage;  
RX MEDLINE=98442651; PubMed=9771708;  
RA Ul Haque M.F., King L.M., Krakow D., Cantor R.M., Rusiniak M.E.,  
RA Swank R.T., Superti-Furga A., Haque S., Abbas H., Ahmad W., Ahmad M.,  
RA Cohn D.H.;  
RT "Mutations in orthologous genes in human spondyloepimetaphyseal  
RT dysplasia and the brachymorphic mouse."  
RL Nat. Genet. 20:157-162(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RA Franzon V.L., Gibson M.A., Hatzinikolas G., Cleary E.G., Woolatt E.,  
RA Sutherland G.R.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RA Fuda H., Shimizu C., Strott C.A.;  
RT "Human bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase:  
RT differential expression of isoforms and effect of polymorphisms on  
RT activity."  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RX MEDLINE=20145452; PubMed=10679223;  
RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,  
RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,  
RA Weinshilboum R.M.;  
RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and  
RT PAPSS2: gene cloning, characterization and chromosomal localization."  
RL Biochem. Biophys. Res. Commun. 268:437-444(2000).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM B).  
RC TISSUE=Liver;  
RX MEDLINE=20026854; PubMed=10559207;  
RA Kurima K., Singh B., Schwartz N.B.;  
RT "Genomic organization of the mouse and human genes encoding the ATP  
RT sulfurylase/adenosine 5'-phosphosulfate kinase isoform SK2."  
RL J. Biol. Chem. 274:33306-33312(1999).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM B).  
RA Venkatachalam K.V., Fuda H., Strott C.A.;  
RT "3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM A).  
RC TISSUE=Colon;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS  
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE  
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE  
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE  
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS  
CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS). ACTIVATED SULFATE DONOR  
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF  
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-  
CC ACTIVATION PATHWAY. MAY HAVE AN IMPORTANT ROLE IN SKELETOGENESIS  
CC DURING POSTNATAL GROWTH (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-





RL Nat. Genet. 20:157-162(1998).

CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS

CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE

CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE

CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE

CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS

CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS): ACTIVATED SULFATE DONOR

CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF

CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-

CC ACTIVATION PATHWAY. MAY HAVE A IMPORTANT ROLE IN SKELETOGENESIS

CC DURING POSTNATAL GROWTH.

CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.

CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

CC phosphoadenylylsulfate.

CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO

CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION

CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, CARTILAGE, SKIN AND BRAIN.

CC -!- DISEASE: DEFECTS IN PAPS2 ARE THE CAUSE OF BRACHYMRPHISM (BM), A

CC AUTOSOMAL RECESSIVE DISEASE, WHICH IS CHARACTERIZED BY ABNORMAL

CC HEPATIC DETOXIFICATION, BLEEDING TIMES AND POSTNATAL GROWTH, SUCH

CC AS HOME-SHAPED SKULL, SHORT THICK TAIL, AND SHORTENED BUT NOT

CC WIDENED LIMBS. THE ABNORMAL POSTNATAL GROWTH HAS BEEN ATTRIBUTED

CC TO UNDERSULFATION OF CARTILAGE PROTEOGLYCANS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE

CC FAMILY.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE

CC ADENYLYLSULFOTRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF052453; AAC40191.1; -

DR EMBL; AF085144; AAC98687.1; -

DR MGD; MGI:1330223; Paps2.

DR InterPro; IPR002891; APS\_Kinase.

DR Pfam; PF01583; APS\_kinase; 1

DR Pfam; PF01747; ATP-sulfurylase; 1

DR ProDom; PD002350; APS\_Kinase; 1

DR ProDom; PD002381; ATP-sulfurylase; 1

DR TIGRFAMs; TIGR00455; apsk; 1

DR ATP-binding; Multigene family; Disease mutation; Polymorphism.

DR ADENYLYLSULFATE KINASE.

DR DOMAIN 1 7211

DR DOMAIN 7212 621

DR NP\_BIND 50 57

DR ACT\_SITE 124 124

DR SITE 517 521

DR VARIANT 79 79

DR VARIANT 109 109

DR CONFLICT 5 5

DR CONFLICT 290 294

DR CONFLICT 621 AA; 70290 MW; DE375F06A79FEAAB CRC64;

DR SEQUENCE 621 AA; 70290 MW; DE375F06A79FEAAB CRC64;

Query Match 13.2%; Score 81; DB 1; Length 621;

Best Local Similarity 100.0%; Pred. No. 5.4e-75;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 LFADAGLVCTSFSPAKDRENARKTHESAGLPFFEIFVDAPLNICESRDVKGLYRRAR 168

DB 110 LFADAGLVCTSFSPAKDRENARKTHESAGLPFFEIFVDAPLNICESRDVKGLYRRAR 169

QY 169 AGEIKGTGIDSDYKPEPTE 189

DB 170 AGEIKGTGIDSDYKPEPTE 190

RESULT 3

PPSI\_CAVPO STANDARD; PRT; 624 AA.

AC 054820;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS

DE synthetase 1) (PAPS1) (Sulfurylase kinase 1) (SK1) (SK 1)

DE [Includes: sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylylate

DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase

DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)

DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-

DE phosphoadenosine-5'-phosphosulfate synthetase)]

GN PAPS1 OR PAPS OR ATPSKI.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NH 2;

RA Venkatachalam K V., Akita H., Strott C.A.;

RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.

CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS

CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE

CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE

CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE

CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS

CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS): ACTIVATED SULFATE DONOR

CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF

CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-

CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED

CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.

CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

CC phosphoadenylylsulfate.

CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO

CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION

CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE

CC FAMILY.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE

CC ADENYLYLSULFOTRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF004875; AAC02266.1; -

DR InterPro; IPR002891; APS\_Kinase.

DR InterPro; IPR002650; ATP-sulfurylase.

DR Pfam; PF01583; APS\_kinase; 1

DR Pfam; PF01747; ATP-sulfurylase; 1

DR ProDom; PD002350; APS\_Kinase; 1

DR ProDom; PD002381; ATP-sulfurylase; 1

DR TIGRFAMs; TIGR00455; apsk; 1

DR Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;

DR ATP-binding; Multigene family.

DR DOMAIN 1 7220

DR DOMAIN 7221 624

DR NP\_BIND 59 66

DR ACT\_SITE 133 133

DR SITE 521 525

DR SEQUENCE 624 AA; 70395 MW; BF7461B4D07E2131 CRC64;

Query Match 6.2%; Score 38; DB 1; Length 624;



DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)  
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-  
DE phosphoadenosine-5'-phosphosulfate synthetase)).  
GN PAPSS1 OR PAPSS OR ASAPK OR ATPSK1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Brain;  
RC MEDLINE=96094345; PubMed=7493984;  
RA Li H., Deyrup A., Mensch J.R. Jr., Domowicz M., Konstantinidis A.K.,  
RA Schwartz N.B.;  
RT "The isolation and characterization of cDNA encoding the mouse  
RT bifunctional ATP sulfonylase-adenosine 5'-phosphosulfate kinase.";  
RL J. Biol. Chem. 270:29453-29459(1995).  
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS  
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE  
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE  
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE  
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS  
CC YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS). ACTIVATED SULFATE DONOR  
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF  
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-  
CC ACTIVATION PATHWAYS. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED  
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.  
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO  
CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION  
CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE NEONATAL BRAIN AND IN  
CC CARTILAGE.  
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE  
CC FAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE  
CC ADENYLYLTRANSFERASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U34883; AAC52328.1; -  
DR MGD: MGI:1330587; Paps1.  
DR InterPro: IPR002891; APS\_kinase.  
DR Pfam: PF01583; APS\_kinase; 1.  
DR Pfam: PF01747; ATP-sulfonylase; 1.  
DR ProDom: PD002350; APS\_kinase; 1.  
DR ProDom: PD002381; ATP-sulfonylase; 1.  
DR TIGRfams: TIGR00455; apsk; 1.  
DR Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;  
KW ATP-binding; Multigene family.  
FT DOMAIN 1 7220 ADENYLYLSULFATE KINASE.  
FT DOMAIN 7221 624 SULFATE ADENYLYLTRANSFERASE.  
FT NP\_BIND 59 66 ATP (POTENTIAL).  
FT ACT\_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT PP-MOTIF.  
FT SITE 521 525 PP-MOTIF.  
FT SEQUENCE 624 AA; 70794 MW; B487EFAF9B78BE3E CRC64;  
Query Match 6.2%; Score 38; DB 1; Length 624;  
Best Local Similarity 100.0%; Pred. No. 1e-30;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 20 YQAHVSRNRKRGVGTGRCGTCVWLTGLSGAGKTT 57  
|||||

Db 30 YQAHVSRNRKRGVGTGRCGTCVWLTGLSGAGKTT 67  
RESULT 6  
KAPS\_EMENI  
ID KAPS\_EMENI STANDARD; PRT; 206 AA.  
AC Q92203;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
DE 5'-phosphosulfate kinase) (ATP adenosine 5'-phosphosulfate 3'-  
DE phosphotransferase).  
GN SO.  
OS Emericella nidulans (Aspergillus nidulans).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eutiales; Trichocomaceae; Emericella.  
OX NCBI\_TaxID=5072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Clarke D.L., Newbert R.W., Turner G.;  
RT "Cloning and characterisation of the APS kinase gene from  
RT Aspergillus nidulans.";  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT  
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.  
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Y08866; CAA70089.1; -  
DR InterPro: IPR002891; APS\_Kinase.  
DR Pfam: PF01583; APS\_Kinase; 1.  
DR ProDom: PD002350; APS\_Kinase; 1.  
DR TIGRfams: TIGR00455; apsk; 1.  
DR Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
KW Phosphorylation.  
FT NP\_BIND 31 38 ATP (POTENTIAL).  
FT ACT\_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
FT SEQUENCE 206 AA; 23028 MW; 6AADD483E2BCA1CD CRC64;  
Query Match 2.4%; Score 15; DB 1; Length 206;  
Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 98 ENIRRIAEVAKLFAD 112  
|||||  
Db 80 ENIRRIAEVAKLFAD 94  
RESULT 7  
KAPS\_PENCH  
ID KAPS\_PENCH STANDARD; PRT; 211 AA.  
AC Q12657;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
DE 5'-phosphosulfate kinase) (ATP adenosine 5'-phosphosulfate 3'-  
DE phosphotransferase).  
OS Penicillium chrysogenum.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.

```

OX NCBI_TaxID=5076;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 24791;
RA Foster B.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39393; AAA1521.1;
DR InterPro; IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT ACT_SITE 107 107 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
FT SQ SEQUENCE 211 AA; 23770 MW; 7DDC4BDA867FE7C2 CRC64;

Query Match 2.48; Score 15; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 ENIRRIAEVAKLFAD 112
DB 82 ENIRRIAEVAKLFAD 96
|||||
|||||

RESULT 8
CYSC_SYNY3
ID CYSC_SYNY3 STANDARD; PRT; 177 AA.
AC P72940;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase).
DE DE PHOSPHOTRANSFERASE.
GN CYSC OR SLR0676.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE

```

```

CC BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90902; BAA16957.1;
DR InterPro; IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 12 19 ATP (BY SIMILARITY).
FT ACT_SITE 86 86 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
FT SQ SEQUENCE 177 AA; 19674 MW; E24D5B5452FCD7D8 CRC64;

Query Match 2.18; Score 13; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 WLTLGSLGAGKTTI 58
DB 9 WLTLGSLGAGKTTI 21
|||||
|||||

RESULT 9
KAP1_ARATH
ID KAP1_ARATH STANDARD; PRT; 276 AA.
AC Q43295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 1, chloroplast precursor (EC 2.7.1.25) (APS
DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
DE phosphosulfate 3'-phosphotransferase).
DE GN AKN1 OR AT2G14750 OR F26C24.11 OR T6B13.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=94325358; PubMed=8049272;
RA Arz H.E., Gisselmann G., Schiffmann S., Schwen J.-D.;
RT "A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis
RT thaliana.";
RL Blochim. Biophys. Acta 1218:447-452(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94345022; PubMed=8066145;
RA Jain A., Leustek T.;
RT "A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
RT thaliana.";
RL Plant Physiol. 105:771-772(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RA Leustek T.;
RT "Analysis of ATP-sulfurylase and APS-kinase genomic DNA sequences
RT in Arabidopsis thaliana.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

```

```

RX MEDLINE-20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75782; CAA53426.1; -
DR EMBL; U50238; AAC50035.1; -
DR EMBL; U59759; AAC50034.1; -
DR EMBL; AC004705; AAC24182.1; -
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
DR Transphorase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Kinase; Cysteine biosynthesis; Chloroplast.
FT TRANSIT 1 37
FT CHAIN 38 276
FT NP_BIND 108 115
FT ACT_SITE 182 182
FT FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 276 AA; 29787 MW; CB698643AA09D811 CRC64;
Query Match 2.0%; Score 12; DB 1; Length 276;
Best Local Similarity 100.0%; pred. No. 0.00029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 76 LDGDNVRHGLNR 87
Db 135 LDGDNVRHGLNR 146
RESULT 10
KAP2_ARATH STANDARD; PRT; 293 AA.
AC 049196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS
DE kinase) (adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
DE phosphosulfate 3'-phosphotransferase).
GS AKN2 OR AT4G39940 OR T517.110.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=cv. Columbia;
RA Schiffmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase
RT (EC2.7.1.25) from Catharanthus roseus and an isoform (aku2) from
RT Arabidopsis.";
RL (In) Plant Gene Register PGR98-116.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA MEDLINE-20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansong W., Brandt P., Griwall L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Giesen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Schaefer M., Grimm M., Loehnert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Farmann B., Grandrath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenat O., Quigley F., Clabault G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefod F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Hejlsman L., Schwarz S., Scholler P., Heber S., Franco P., Biele C.,
RA Freshman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; AF043351; AAC39520.1; -
DR EMBL; AF035708; CAB38907.1; -
DR EMBL; AL161596; CAB80657.1; -
DR InterPro; IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
Phosphorylation; Transient peptide; Chloroplast.
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN ? 293 ADENYLSULFATE KINASE 2.
FT NP_BIND 142 149 ATP (POTENTIAL).
FT ACT_SITE 196 196 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
SIMILARITY).
SQ SEQUENCE 293 AA; 31977 MW; 71FAD6E9B026886 CRC64;

Query Match 2.0%; Score 12; DB 1; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87
|||||
DB 149 LDGDNVRHGLNR 160

RESULT 11
CYSN_XYLFA STANDARD; PRT; 623 AA.
AC Q9PD78;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CysN/cysc bifunctional enzyme [includes: Sulfate adenylyltransferase
DE subunit 1 (EC 2.7.7.4) (Sulfate adenylate transferase) (SAT) (ATP-
DE sulfurylase large subunit); Adenylylsulfate kinase (EC 2.7.1.25) (APS
DE kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase)].
GN CYSNC OR XF1501.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
```

```

CC -1- FUNCTION: ATP SULFURYLASE MAY BE THE ETPASE, REGULATING ATP
CC SULFURYLASE ACTIVITY (BY SIMILARITY).
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + sulfate -> diphosphate + adenylylsulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate -> ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- PATHWAY: FIRST AND SECOND STEPS IN THE SULFATE ACTIVATION PATHWAY.
CC THESE REACTIONS OCCUR EARLY IN THE REDUCTIVE BRANCH OF THE
CC CYSTEINE BIOSYNTHETIC PATHWAY.
CC -1- SUBUNIT: HETERODIMER COMPOSED OF CYSN, THE SMALLER SUBUNIT, AND
CC CYSN (BY SIMILARITY).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AE003980; AAF84310.1; ALT_INIT.
CC HSSP; P02990; 1ETU.
CC InterPro; IPR002891; APS_Kinase.
CC InterPro; IPR004161; EFTU_D2.
CC InterPro; IPR000795; EF_GTPbind.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF01583; APS_Kinase; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC PRINTS; PR00315; ELONGATNFCT.
CC ProDom; PD002350; APS_Kinase; 1.
CC TIGRFAMs; TIGR00455; apsk; 1.
CC PROSITE; PS00301; EFATOR_GTP; 1.
CC Cysteine biosynthesis; Transferase; Nucleotidyltransferase;
CC GTP-binding; Kinase; ATP-binding; Multifunctional enzyme;
CC Complete proteome.
CC FT DOMAIN 1 450 SULFATE ADENYLYL TRANSFERASE.
CC FT DOMAIN 451 623 ADENYLYLSULFATE KINASE.
CC FT NP_BIND 23 30 GTP (BY SIMILARITY).
CC FT NP_BIND 102 106 GTP (BY SIMILARITY).
CC FT NP_BIND 157 160 GTP (BY SIMILARITY).
CC FT NP_BIND 459 466 ATP (POTENTIAL).
CC FT ACT_SITE 533 533 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
CC SIMILARITY).
CC SQ SEQUENCE 623 AA; 68277 MW; C20730A365B28E94 CRC64;

Query Match 2.0%; Score 12; DB 1; Length 623;
Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87
|||||
DB 486 LDGDNVRHGLNR 497

RESULT 12
NODQ_RHTR STANDARD; PRT; 632 AA.
AC P52978;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylate
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase)].
GN NODQ.
OS Rhizobium tropici.
```

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=398;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CFN 299;  
 RX MEDLINE=96303335; PubMed=8755625;  
 RA Laermans T., Caluwaerts I., Verreth C., Rogel M.A.,  
 RT "Isolation and characterization of Rhizobium tropici Nod factor  
 RT sulfation genes.";  
 RL Mol. Plant Microbe Interact. 9:492-500(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CIAT899;  
 RX MEDLINE=97002746; PubMed=8850086;  
 RA Folch-Mallol J.L., Marroqui S., Sousa C., Manyani H., Lopez-Lara I.M.,  
 RA van der Drift K.M.G.M., Haverkamp J., Quinto C., Gil-Serrano A.,  
 RA Thomas-Oates J., Spaink H.P., Megias M.;  
 RT "Characterization of Rhizobium tropici CIAT899 nodulation factors: the  
 RT role of nodH and nodP genes in their sulfation.";  
 RL Mol. Plant Microbe Interact. 9:151-163(1996).  
 CC -!- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO  
 CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP  
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).  
 CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE  
 CC (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
 CC phosphoadenylylsulfate.  
 CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES. NODP AND NODO, MAY BE  
 CC PHYSICALLY ASSOCIATED.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING  
 CC ELONGATION FACTOR FAMILY. CYSN/NODO SUBFAMILY.  
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U47272; AAB08984.1; -;  
 DR EMBL; X87608; CAA60914.1; -;  
 DR InterPro; IPR002891; APS\_kinase.  
 DR InterPro; IPR004161; EFTU\_D2.  
 DR InterPro; IPR000795; EF\_GTPbind.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.  
 DR TIGRFAMs; TIGR00455; apsk; 1.  
 DR PROSITE; PS00301; EFAC\_TOR\_GTP; 1.  
 KW Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;  
 KW Kinase; ATP-binding; Multifunctional enzyme.  
 FT DOMAIN 1 457 SULFATE ADENYLYL TRANSFERASE.  
 FT NP\_BIND 31 38 GTP (BY SIMILARITY).  
 FT NP\_BIND 109 113 GTP (BY SIMILARITY).  
 FT NP\_BIND 164 167 GTP (BY SIMILARITY).  
 FT NP\_BIND 466 473 ATP (POTENTIAL).  
 FT ACT\_SITE 540 540 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 632 AA; 69814 MW; 8FA3DF0CC7C604B6 CRC64;  
 Query Match 2.0%; Score 12; DB 1; Length 632;  
 Best Local Similarity 100.0%; Pred. No. 0.00059;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 76 LDGDNVRHGLNR 87

DB 493 LDGDNVRHGLNR 504  
 RESULT 13  
 NODO\_RHSB  
 ID NODO\_RHSB STANDARD; PRT; 633 AA.  
 AC 007309;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate  
 DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylyl  
 DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate  
 DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-  
 DE phosphotransferase)].  
 GN NODO.  
 OS Rhizobium sp. (strain BR816).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=48291;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98083747; PubMed=9421916;  
 RA Laermans T., Coolsaet N., Verreth C., Snoeck C., Hellings N.,  
 RA Vanderleyden J., Martinez-Romero E.;  
 RT "Functional redundancy of genes for sulphate activation enzymes in  
 RT Rhizobium sp. BR816.";  
 RL Microbiology 143:3933-3942(1997).  
 CC -!- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO  
 CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP  
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).  
 CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE  
 CC (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
 CC phosphoadenylylsulfate.  
 CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODO, MAY BE  
 CC PHYSICALLY ASSOCIATED (POTENTIAL).  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING  
 CC ELONGATION FACTOR FAMILY. CYSN/NODO SUBFAMILY.  
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U59507; AAB95249.1; -;  
 DR InterPro; IPR002891; APS\_kinase.  
 DR InterPro; IPR004161; EFTU\_D2.  
 DR InterPro; IPR000795; EF\_GTPbind.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR TIGRFAMs; TIGR00455; apsk; 1.  
 DR PROSITE; PS00301; EFAC\_TOR\_GTP; 1.  
 KW Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;  
 KW Kinase; ATP-binding; Multifunctional enzyme.  
 FT DOMAIN 1 458 SULFATE ADENYLYL TRANSFERASE.  
 FT NP\_BIND 31 38 GTP (BY SIMILARITY).  
 FT NP\_BIND 110 114 GTP (BY SIMILARITY).  
 FT NP\_BIND 165 168 GTP (BY SIMILARITY).  
 FT NP\_BIND 467 474 ATP (POTENTIAL).  
 FT ACT\_SITE 541 541 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY

FT SQ SEQUENCE 633 AA; 69279 MW; 696BCA683DB4B3111 CRC64; SIMILARITY).

Query Match 2.0%; Score 12; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 0.00059;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87  
|||||  
DB 494 LDGDNVRHGLNR 505

RESULT 14  
ID NODQ\_RHIME STANDARD; PRT; 641 AA.  
AC P13442;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate  
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylylate  
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate  
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-  
DE phosphotransferase)].  
DE NODQ OR RA0469 OR SMA0857.  
GN Rhizobium meliloti (Sinorhizobium meliloti).  
OS Plasmid pSymA (megaplasmid 1).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RCR2011 / S047;  
RX MEDLINE=89313304; PubMed=2546009;  
RA Cervantes E., Sharma S.B., Mailliet F., Vasse J., Truchet G.,  
RA Rosenberg C.;  
RT "The Rhizobium meliloti host range nodQ gene encodes a protein which  
RT shares homology with translation elongation and initiation factors.";  
RL Mol. Microbiol. 3:745-755(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92288403; PubMed=2520820;  
RA Schwedock J., Long S.R.;  
RT "Nucleotide sequence and protein products of two new nodulation genes  
RT of Rhizobium meliloti, nodP and nodQ.";  
RL Mol. Plant Microbe Interact. 2:181-194(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=1021;  
RX MEDLINE=21396509; PubMed=11481432;  
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
RT "Nucleotide sequence and predicted functions of the entire  
RT Sinorhizobium meliloti pSymA megaplasmid.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9889(2001).  
CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO  
CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP  
CC SULFURYLASE ACTIVITY (BY SIMILARITY).  
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE  
CC (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC -1- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODQ, MAY BE  
CC PHYSICALLY ASSOCIATED (POTENTIAL).  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING  
CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE  
CC FAMILY.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC DR EMBL; X14809; CAA32914.1; -;  
CC DR EMBL; M68858; AAA26343.1; -;  
CC DR EMBL; AE007237; AKK65127.1; -;  
CC DR PIR; S14899; ZZZRNO.  
CC DR InterPro: IPR002891; APS\_kinase.  
CC DR InterPro: IPR004161; EFTU\_D2.  
CC DR InterPro: IPR000795; EF\_GTPbind.  
CC DR InterPro: IPR005225; Small\_GTP.  
CC DR Pfam; PF00009; GTP\_EFTU; 1.  
CC DR Pfam; PF01583; APS\_kinase; 1.  
CC DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
CC DR ProDom; PD002350; APS\_kinase; 1.  
CC DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
CC DR TIGRFAMs; TIGR00455; aps; 1.  
CC DR PROSITE; PS00301; EFATOR\_GTP; 1.  
CC Plasmid; Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;  
CC Kinase; ATP-binding; Multifunctional enzyme; Complete proteome.  
KW DOMAIN 1 458 SULFATE ADENYLYL TRANSFERASE.  
FT NP\_BIND 31 38 ADENYLYLSULFATE KINASE.  
FT NP\_BIND 110 114 GTP (BY SIMILARITY).  
FT NP\_BIND 165 168 GTP (BY SIMILARITY).  
FT NP\_BIND 467 474 GTP (BY SIMILARITY).  
FT ACT\_SITE 524 524 ATP (POTENTIAL).  
FT ACT\_SITE 524 524 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
SQ SEQUENCE 641 AA; 70614 MW; 1E1261F04ED33A93 CRC64;  
Query Match 2.0%; Score 12; DB 1; Length 641;  
Best Local Similarity 100.0%; Pred. No. 0.0006;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87  
|||||  
DB 494 LDGDNVRHGLNR 505

RESULT 15  
ID NODQ\_RHIS3 STANDARD; PRT; 646 AA.  
AC P72339;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate  
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylylate  
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate  
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-  
DE phosphotransferase)].  
DE NODQ.  
GN Rhizobium sp. (strain N33).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=103798;  
RN [1]  
RP SEQUENCE FROM N.A.  
RL Cloutier J.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO  
CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP  
CC SULFURYLASE ACTIVITY (BY SIMILARITY).  
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE  
CC (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC



```

CC CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODQ. MAY BE
CC CC PHYSICALLY ASSOCIATED (POTENTIAL).
CC CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
CC CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC CC FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U53327; AAB16902.1; -.
CC CC HSSP; P02990; 1ETU.
CC CC InterPro; IPR002891; APS_kinase.
CC CC InterPro; IPR004161; EFTU_D2.
CC CC InterPro; IPR000795; EF-GTPbind.
CC CC Pfam; PF00009; GTP-EFTU; 1.
CC CC Pfam; PF01583; APS_kinase; 1.
CC CC Pfam; PF03144; GTP-EFTU_D2; 1.
CC CC ProDom; PD002350; APS_kinase; 1.
CC CC TIGRfam; TIGR00455; apsk; 1.
CC CC PROSITE; PS00301; EFACFACOR_GTP; 1.
CC CC Modulation; Transferase; Nucleotidyltransferase; Gtp-binding;
CC CC Kinase; ATP-binding; Multifunctional enzyme.
CC CC FT DOMAIN 1 461 SULFATE ADENYLYL TRANSFERASE.
CC CC FT DOMAIN 462 646 ADENYLYLSULFATE KINASE.
CC CC FT NP_BIND 34 41 GTP (BY SIMILARITY).
CC CC FT NP_BIND 113 117 GTP (BY SIMILARITY).
CC CC FT NP_BIND 168 171 GTP (BY SIMILARITY).
CC CC FT NP_BIND 470 477 ATP (POTENTIAL).
CC CC FT ACT_SITE 544 544 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
CC CC SIMILARITY).
CC CC SEQUENCE 646 AA: 71488 MW: 1DBEE1DA257FE128 CRC64;
CC CC
CC CC Query Match 2.0%; Score 12; DB 1; Length 646;
CC CC Best Local Similarity 100.0%; Pred. No. 0.0006;
CC CC Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC CC
CC CC QY 76 LDGDNVVRHGLNR 87
CC CC 111111111111
CC CC Db 497 LDGDNVVRHGLNR 508

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:26:25 ; Search time 87 Seconds  
(without alignments)  
1456.540 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 615  
Sequence: 1 MSGIKKQKTENQOKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 10

Total number of hits satisfying chosen parameters: 51

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertbrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID        | Description         |
|------------|-------|-------------|--------|-----------|---------------------|
| 1          | 142   | 23.1        | 265    | 4 Q9UIR2  | Q9UIR2 homo sapien  |
| 2          | 81    | 13.2        | 621    | 11 Q9QYSO | Q9QYSO mus musculus |
| 3          | 58    | 9.4         | 620    | 11 Q9JK86 | Q9JK86 cavia porce  |
| 4          | 38    | 6.2         | 624    | 4 Q9UE98  | Q9UE98 homo sapien  |
| 5          | 38    | 6.2         | 624    | 4 Q9E1P9  | Q9E1P9 homo sapien  |
| 6          | 38    | 6.2         | 624    | 4 Q9GTF4  | Q9GTF4 homo sapien  |
| 7          | 37    | 6.0         | 613    | 13 Q90XY2 | Q90XY2 fugu rubrip  |
| 8          | 26    | 4.2         | 618    | 5 Q9NDP8  | Q9NDP8 ciona intes  |
| 9          | 22    | 3.6         | 424    | 10 Q43170 | Q43170 solanum tub  |
| 10         | 22    | 3.6         | 458    | 10 Q8W1X3 | Q8W1X3 allium cepa  |
| 11         | 22    | 3.6         | 461    | 10 Q9SDP4 | Q9SDP4 allium cepa  |
| 12         | 22    | 3.6         | 461    | 10 Q8SAG1 | Q8SAG1 glycine max  |
| 13         | 22    | 3.6         | 463    | 10 Q42519 | Q42519 arabidopsis  |
| 14         | 22    | 3.6         | 463    | 10 Q9SE02 | Q9SE02 arabidopsis  |
| 15         | 22    | 3.6         | 463    | 10 Q43183 | Q43183 solanum tub  |
| 16         | 22    | 3.6         | 463    | 10 Q9L1K9 | Q9L1K9 arabidopsis  |

## ALIGNMENTS

## RESULT 1

| ID | Q9UIR2  | PRELIMINARY; | PRT; | 265 AA. |
|----|---|--------------|------|---------|
| AC | Q9UIR2;   |              |      |         |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created)                                |              |      |         |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update)                   |              |      |         |
| DT | 01-MAR-2002 (TREMBLrel. 20, Last annotation update)                 |              |      |         |
| DE | ATP sulfurylase/APS kinase isoform SK2 (EC 2.7.7.4) (Fragment).     |              |      |         |
| OS | Homo sapiens (human).   |              |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |              |      |         |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.          |              |      |         |
| OX | NCBI_TaxID=9606;  |              |      |         |
| RN | [1]   |              |      |         |
| RP | SEQUENCE FROM N.A.  |              |      |         |
| RX | MEDLINE=98337975; PubMed=9671738;                                   |              |      |         |
| RA | Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr., |              |      |         |
| RA | Deyrup A., Schwartz N.B.;   |              |      |         |
| RT | "A member of a family of sulfate-activating enzymes causes murine   |              |      |         |
| RT | brachymorphism.";   |              |      |         |
| RN | [2]   |              |      |         |
| RP | Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).                   |              |      |         |
| RX | SEQUENCE FROM N.A.  |              |      |         |
| RA | MEDLINE=20026854; PubMed=10559207;                                  |              |      |         |
| RA | Kurima K., Singh B., Schwartz N.B.;                                 |              |      |         |
| RT | "Genomic organization of the mouse and human genes encoding the ATP |              |      |         |
| RT | Sulfurylase/Adenosine 5'-phosphosulfate kinase isoform SK2.";       |              |      |         |
| RL | J. Biol. Chem. 274:33306-33312(1999).                               |              |      |         |
| DR | EMBL; AF173364; AAF12780.1;   |              |      |         |
| DR | EMBL; AF173361; AAF12780.1; JOINED.                                 |              |      |         |
| DR | EMBL; AF173362; AAF12780.1; JOINED.                                 |              |      |         |
| DR | EMBL; AF173363; AAF12780.1; JOINED.                                 |              |      |         |
| DR | InterPro: IPR002650; ATP-sulfurylase.                               |              |      |         |
| DR | Pfam: PF01747; ATP-sulfurylase; 1.                                  |              |      |         |

Q92wm0 oryza sativ  
Q48888 zea mays (m  
Q27128 urechis cau  
Q96349 brassica na  
Q92n28 brassica ju  
Q42520 arabidopsis  
Q96530 arabidopsis  
Q23324 arabidopsis  
Q957d8 arabidopsis  
Q92n29 brassica ju  
Q43870 arabidopsis  
Q8rwj3 arabidopsis  
Q96541 brassica ol  
Q96b10 brassica ol  
Q961a8 drosophila  
Q9vw48 drosophila  
Q95p41 aedes aegypt  
Q22501 caenorhabdl  
Q9p7g9 schizosacch  
Q9se92 zea mays (m  
Q92vh6 rhizobium m  
Q985q4 rhizobium l  
Q9sqr9 arabidopsis  
Q9sqr7 arabidopsis  
Q81351 enteromorph  
Q9fjx1 arabidopsis  
Q9a882 caulobacter  
Q8yex3 bruceella me  
Q97mt8 clostridium  
Q9eyl1 klebsiella  
Q8xf34 salmonella  
Q9hgf8 saccharomyc  
Q9c3y6 saccharomyc  
Q8zbp3 versinia pe  
Q8rell fusobacteri

DR ProDom: PD002381; ATP-sulfonylase; 1.  
KW Kinase; Nucleotidyltransferase; Transferase.  
FT NON\_TER 1  
SQ SEQUENCE 265 AA; 30142 MW; D75A2374E40C334F CRC64;

Query Match 23.18; Score 142; DB 4; Length 265;  
Best Local Similarity 100.0%; Pred. No. 1.5e-145;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 427 RLLERGYKHPVLLHPLGLGWTGKDDVPLDWRMKQHAHVLEEGVLDPKSTIVAIFFSPML 486  
|||||  
Db 70 RLLERGYKHPVLLHPLGLGWTGKDDVPLDWRMKQHAHVLEEGVLDPKSTIVAIFFSPML 129  
|||||

Qy 487 YAGTEVQWCHRSRMIAAGFYIVGRDPAGMPHPETKKDLYEPTHGKVLSPMAPGLTSVE 546  
|||||  
Db 130 YAGTEVQWCHRSRMIAAGFYIVGRDPAGMPHPETKKDLYEPTHGKVLSPMAPGLTSVE 189  
|||||

Qy 547 IIPRVAAYNKKAKAMDFYDPA 568  
|||||  
Db 190 IIPRVAAYNKKAKAMDFYDPA 211  
|||||

RESULT 2  
Q9QYS0 PRELIMINARY; PRT; 621 AA.  
ID Q9QYS0  
AC Q9QYS0  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE ATP sulfonylase/APS kinase isoform SK2 (EC 2.7.7.4).  
GN PAPSS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE=20026854; PubMed=10559207;  
RA Kurima K., Singh B., Schwartz N.B.;  
RA Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr.,  
RA Deyrup A., Schwartz N.B.;  
RT "A member of a family of sulfate-activating enzymes causes murine  
RT brachycephalism."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RX MEDLINE=20026854; PubMed=10559207;  
RA Kurima K., Singh B., Schwartz N.B.;  
RT "Genomic organization of the mouse and human genes encoding the ATP  
RT Sulfonylase/Adenosine 5'-phosphosulfate kinase isoform SK2.";  
RL J. Biol. Chem. 274:33306-33312(1999).  
DR EMBL; AF172866; AAF12760.1; JOINED.  
DR EMBL; AF172867; AAF12760.1; JOINED.  
DR EMBL; AF172868; AAF12760.1; JOINED.  
DR EMBL; AF172869; AAF12760.1; JOINED.  
DR EMBL; AF172870; AAF12760.1; JOINED.  
DR EMBL; AF172871; AAF12760.1; JOINED.  
DR EMBL; AF172872; AAF12760.1; JOINED.  
DR EMBL; AF172873; AAF12760.1; JOINED.  
DR EMBL; AF172874; AAF12760.1; JOINED.  
DR EMBL; AF172875; AAF12760.1; JOINED.  
DR MGI; MGI.1330223; Papss2.  
DR InterPro; IPR002891; APS\_kinase.  
DR InterPro; IPR002650; ATP-sulfonylase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfonylase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR ProDom; PD002381; ATP-sulfonylase; 1.  
DR TIGRFAMS; TIGR00455; apsk; 1.  
DR Kinase; Nucleotidyltransferase; Transferase.  
SQ SEQUENCE 621 AA; 70408 MW; 019519DC8E239F68 CRC64;

Query Match 13.28; Score 81; DB 11; Length 621;  
Best Local Similarity 100.0%; Pred. No. 9.2e-79;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 LFADAGLVCTISFISPPAKDRENARKTHESAGLFFFEIFVDAPLNICESRDVKGLYKRR 168  
|||||  
Db 110 LFADAGLVCTISFISPPAKDRENARKTHESAGLFFFEIFVDAPLNICESRDVKGLYKRR 169  
|||||

Qy 169 AGEIKGFTGDSYKPKETPE 189  
|||||  
Db 170 AGEIKGFTGDSYKPKETPE 190  
|||||

RESULT 3  
Q9JK86 PRELIMINARY; PRT; 620 AA.  
ID Q9JK86  
AC Q9JK86  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Adenosine 5'-phosphosulfate kinase/ATP sulfonylase 2.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fuda H., Strott C.A.;  
RT "Guinea pig bifunctional adenosine 5'-phosphosulfate kinase/ATP  
RT sulfonylase 2, gpAPS synthase 2.";  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF251798; AAF70194.1; JOINED.  
DR InterPro; IPR002891; APS\_kinase.  
DR InterPro; IPR002650; ATP-sulfonylase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfonylase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR ProDom; PD002381; ATP-sulfonylase; 1.  
DR TIGRFAMS; TIGR00455; apsk; 1.  
DR Kinase.  
SQ SEQUENCE 620 AA; 70031 MW; 50E187D16335985A CRC64;

Query Match 9.48; Score 58; DB 11; Length 620;  
Best Local Similarity 100.0%; Pred. No. 1e-53;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 500 RMIAGANFYIVGRDPAGMPHPETKKDLYEPTHGKVLSPMAPGLTSVEIIPRVAAYNK 557  
|||||  
Db 505 RMIAGANFYIVGRDPAGMPHPETKKDLYEPTHGKVLSPMAPGLTSVEIIPRVAAYNK 562  
|||||

RESULT 4  
Q9UE98 PRELIMINARY; PRT; 624 AA.  
ID Q9UE98  
AC Q9UE98  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE ATP sulfonylase/APS kinase.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Deyrup A.T.;  
RT "Human ATP sulfonylase/APS kinase.";  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF016496; AAD09325.1; JOINED.  
DR InterPro; IPR002891; APS\_kinase.  
DR InterPro; IPR002650; ATP-sulfonylase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfonylase; 1.

DR ProDom; PD002350; APS\_kinase; 1.  
 DR ProDom; PD002381; ATP-sulphurylase; 1.  
 DR TIGRFAMS; TIGR00455; apsk; 1.  
 KW Kinase.  
 SQ SEQUENCE 624 AA; 70881 MW; 0008DEC57B6F35BC CRC64;

Query Match 6.2%; Score 38; DB 4; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57  
 |||||||  
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67

RESULT 5  
 Q9P1P9  
 ID Q9P1P9 PRELIMINARY; PRT; 624 AA.  
 AC Q9P1P9;  
 DT 01-OCT-2000 (TremBLrel. 15, Created)  
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)  
 DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.  
 GN PAPSS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20145452; PubMed=10679223;  
 RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,  
 RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,  
 RA Weinshilboum R.M.;  
 RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and  
 RT PAPSS2: gene cloning, characterization and chromosomal localization.";  
 RL Biochem. Biophys. Res. Commun. 268:437-444(2000).  
 DR EMBL; AF015227; AAF40236.1; JOINED.  
 DR InterPro; IPR002650; APS\_kinase.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.  
 DR ProDom; PD002381; ATP-sulphurylase; 1.  
 DR TIGRFAMS; TIGR00455; apsk; 1.  
 SQ SEQUENCE 624 AA; 70859 MW; 60DC9B943E7B75ED CRC64;

Query Match 6.2%; Score 38; DB 4; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57  
 |||||||  
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67

RESULT 6  
 Q96TF4  
 ID Q96TF4 PRELIMINARY; PRT; 624 AA.  
 AC Q96TF4; Q96FBI;  
 DT 01-DEC-2001 (TremBLrel. 19, Created)  
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)  
 DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.  
 GN PAPSS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20145452; PubMed=10679223;

RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,  
 RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,  
 RA Weinshilboum R.M.;  
 RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and  
 RT PAPSS2: gene cloning, characterization and chromosomal localization.";  
 RL Biochem. Biophys. Res. Commun. 268:437-444(2000).  
 RN [2]  
 RC SEQUENCE OF 22-624 FROM N.A.  
 RC TISSUE=EYE;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF097721; AAF40235.1; JOINED.  
 DR EMBL; AF097710; AAF40235.1; JOINED.  
 DR EMBL; AF097711; AAF40235.1; JOINED.  
 DR EMBL; AF097712; AAF40235.1; JOINED.  
 DR EMBL; AF097713; AAF40235.1; JOINED.  
 DR EMBL; AF097714; AAF40235.1; JOINED.  
 DR EMBL; AF097715; AAF40235.1; JOINED.  
 DR EMBL; AF097716; AAF40235.1; JOINED.  
 DR EMBL; AF097717; AAF40235.1; JOINED.  
 DR EMBL; AF097718; AAF40235.1; JOINED.  
 DR EMBL; AF097719; AAF40235.1; JOINED.  
 DR EMBL; AF097720; AAF40235.1; JOINED.  
 DR EMBL; BC011392; AAH11392.1; JOINED.  
 DR InterPro; IPR002891; APS\_kinase.  
 DR InterPro; IPR002650; ATP-sulphurylase.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR Pfam; PF01747; ATP-sulphurylase; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.  
 DR ProDom; PD002381; ATP-sulphurylase; 1.  
 DR TIGRFAMS; TIGR00455; apsk; 1.  
 SQ SEQUENCE 624 AA; 70833 MW; A3DC9B943E68CDD6 CRC64;

Query Match 6.2%; Score 38; DB 4; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57  
 |||||||  
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67

RESULT 7  
 Q90XY2  
 ID Q90XY2 PRELIMINARY; PRT; 613 AA.  
 AC Q90XY2;  
 DT 01-DEC-2001 (TremBLrel. 19, Created)  
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)  
 DE 3'-phosphoadenosine 5'-phosphosulfate synthase 2.  
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Takifugu.  
 OX NCBI\_TaxID=31033;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RX MEDLINE=21455682; PubMed=11571655;  
 RA Yu W.P., Pallen C.J., Tay A., Jirik F.R., Brenner S., Tan Y.H.,  
 RA Venkatesh B.;  
 RT "Conserved synteny between the Fugu and human PTEN locus and the  
 RT evolutionary conservation of vertebrate PTEN function.";  
 RL Oncogene 20:5554-5561(2001).  
 DR EMBL; AF325922; AAL08416.1; JOINED.  
 DR InterPro; IPR002891; APS\_kinase.  
 DR InterPro; IPR002650; ATP-sulphurylase.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR Pfam; PF01747; ATP-sulphurylase; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.  
 DR ProDom; PD002381; ATP-sulphurylase; 1.  
 DR TIGRFAMS; TIGR00455; apsk; 1.  
 SQ SEQUENCE 613 AA; 68948 MW; 28E50148377C4169 CRC64;

Query Match 6.0%; Score 37; DB 13; Length 613;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-31;  
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 PVLLHPLGGTWDVPLDWRMKQHAHVLEEGVLPD 473  
 Db 436 PVLLHPLGGTWDVPLDWRMKQHAHVLEEGVLPD 472

## RESULT 8

Q9NDP8 ID Q9NDP8 PRELIMINARY; PRT; 618 AA.  
 AC Q9NDP8; 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ATP sulfurylase/APS kinase.  
 GN CI-ASAK.  
 OS Clona intestinalis.  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 OC Phlebobranchia; Clonidae; Clona.  
 OX NCBI\_TaxID=7719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,  
 RA Satou N.;  
 RT "Characterization of Brachyury downstream notochord genes in the Clona  
 RL intestinalis embryo.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RM EMBL: AB036852; BAB00629.1; -  
 DR InterPro: IPR002891; APS\_kinase.  
 DR InterPro: IPR002650; ATP-sulfurylase.  
 DR Pfam: PF01583; APS\_kinase; 1.  
 DR ProDom: PD002350; APS\_kinase; 1.  
 DR ProDom: PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMs: TIGR00455; apsk; 1.  
 KW Kinase.  
 SQ SEQUENCE 618 AA; 69559 MW; 75223CCB02BC397 CRC64;

Query Match 4.2%; Score 26; DB 5; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-19;  
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MNADAVFAFQLRNPVHNGHALLMQDT 425  
 Db 403 MNADAVFAFQLRNPVHNGHALLMQDT 428

## RESULT 9

Q43170 ID Q43170 PRELIMINARY; PRT; 424 AA.  
 AC Q43170;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Sulfate adenylyltransferase (EC 2.7.7.4).  
 GN STMET3-1.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DESIREE; TISSUE=LEAF;  
 RX MEDLINE=95004649; PubMed=7920699;  
 RA Klonous D., Hoergen R., Willmitzer L., Riesmeier J.W.;  
 RT "Isolation and characterization of two cDNA clones encoding ATP-  
 RT sulfurylase from Potato by complementation of a yeast mutant.";  
 RL Plant J. 6:105-112(1994).  
 DR EMBL: X75041; CAA52953.1; -

DR InterPro: IPR002650; ATP-sulfurylase.  
 DR Pfam: PF01747; ATP-sulfurylase; 1.  
 DR ProDom: PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMs: TIGR00339; sopt; 1.  
 KW Nucleotidyltransferase; Transferase.  
 SQ SEQUENCE 424 AA; 47519 MW; E82A27DC11350ABC CRC64;

Query Match 3.6%; Score 22; DB 10; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422  
 Db 200 NADAVFAFQLRNPVHNGHALLM 221

## RESULT 10

Q8W1X3 ID Q8W1X3 PRELIMINARY; PRT; 458 AA.  
 AC Q8W1X3;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ATP-sulfurylase.  
 OS Allium cepa (Onion).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;  
 OC Allium.  
 OX NCBI\_TaxID=4679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. PUKEKOHE LONG KEEPER; TISSUE=ROOT;  
 RA Pither-Joyce M.D., McCallum J.A.;  
 RT "Isolation of an ATP-sulfurylase cDNA clone from Allium cepa.";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RM EMBL: AF403295; AAL61615.1; -  
 DR InterPro: IPR002650; ATP-sulfurylase.  
 DR Pfam: PF01747; ATP-sulfurylase; 1.  
 DR ProDom: PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMs: TIGR00339; sopt; 1.  
 SQ SEQUENCE 458 AA; 51360 MW; F7A43A53F3A7F04C CRC64;

Query Match 3.6%; Score 22; DB 10; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422  
 Db 233 NADAVFAFQLRNPVHNGHALLM 254

## RESULT 11

Q9SDP4 ID Q9SDP4 PRELIMINARY; PRT; 461 AA.  
 AC Q9SDP4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ATP-sulfurylase.  
 OS Allium cepa (Onion).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;  
 OC Allium.  
 OX NCBI\_TaxID=4679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. PUKEKOHE LONG KEEPER; TISSUE=ROOT;  
 RA Pither-Joyce M.D., McCallum J.A.;  
 RT "Isolation of an ATP-sulfurylase cDNA clone from Allium cepa.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RM EMBL: AF212154; AAF18998.1; -  
 DR InterPro: IPR002650; ATP-sulfurylase.  
 DR Pfam: PF01747; ATP-sulfurylase; 1.

DR ProDom; PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMS; TIGR00339; sopt; 1.  
 SQ SEQUENCE 461 AA; 51807 MW; 04083BE879CBDE71 CRC64;

Query Match 3.6%; Score 22; DB 10; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVAFQLRNPVNHGHALLM 422  
 |||||  
 Db 236 NADAVAFQLRNPVNHGHALLM 257

## RESULT 12

Q8SAG1 PRELIMINARY; PRT; 461 AA.

AC Q8SAG1;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 OS ATP sulfurylase.  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=SEEDLING;  
 RA Phartival P., Krishnan H.B.;

RT "Nucleotide sequence of a soybean cDNA encoding adenosine triphosphate  
 sulfurylase.";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF452454; AAL74418.1; -;  
 SQ SEQUENCE 461 AA; 51700 MW; 3504272C93923100 CRC64;

Query Match 3.6%; Score 22; DB 10; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVAFQLRNPVNHGHALLM 422  
 |||||  
 Db 238 NADAVAFQLRNPVNHGHALLM 259

## RESULT 13

Q42519 PRELIMINARY; PRT; 463 AA.

AC Q42519;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE ATP sulfurylase precursor (EC 2.7.7.4).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=9433677; PubMed=8058839;  
 RA Leustek T., Murillo M., Cervantes M.;  
 RT "Cloning of a cDNA encoding ATP sulfurylase from Arabidopsis thaliana  
 by functional expression in Saccharomyces cerevisiae.";  
 RL Plant Physiol. 105:897-902(1994).

DR EMBL; U05218; AAA21570.1; -;  
 DR InterPro; IPR002650; ATP-sulfurylase.  
 DR Pfam; PF01747; ATP-sulfurylase; 1.  
 DR ProDom; PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMS; TIGR00339; sopt; 1.  
 KW Chloroplast; Nucleotidyltransferase; Transferase; Transit peptide.  
 FT TRANSIT 1 48 CHLOROPLAST.  
 FT CHAIN 49 463 ATP SULFURYLASE.

SQ SEQUENCE 463 AA; 51372 MW; 9A45FAA5133A17FF CRC64;

Query Match 3.6%; Score 22; DB 10; Length 463;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVAFQLRNPVNHGHALLM 422  
 |||||  
 Db 239 NADAVAFQLRNPVNHGHALLM 260

## RESULT 14

Q9SE02 PRELIMINARY; PRT; 463 AA.

AC Q9SE02;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE ATP sulfurylase (EC 2.7.7.4).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20267847; PubMed=10806350;  
 RA Hatzfeld Y., Lee S., Lee M., Leustek T., Saito K.;  
 RT "Functional characterization of a gene encoding a fourth ATP  
 sulfurylase isoform from Arabidopsis thaliana.";  
 RL Gene 248:51-58(2000).  
 DR EMBL; AF198964; AAF19185.1; -;  
 DR InterPro; IPR002650; ATP-sulfurylase.  
 DR Pfam; PF01747; ATP-sulfurylase; 1.  
 DR ProDom; PD002381; ATP-sulfurylase; 1.  
 DR TIGRFAMS; TIGR00339; sopt; 1.  
 DR Nucleotidyltransferase; Transferase.  
 KW SEQUENCE 463 AA; 51444 MW; 18EC241A3B46CB3D CRC64;

Query Match 3.6%; Score 22; DB 10; Length 463;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVAFQLRNPVNHGHALLM 422  
 |||||  
 Db 239 NADAVAFQLRNPVNHGHALLM 260

## RESULT 15

Q43183 PRELIMINARY; PRT; 463 AA.

AC Q43183;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Sulfate adenylyltransferase (EC 2.7.7.4).  
 GN MET3-2.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DESIREE; TISSUE=LEAF;  
 RX MEDLINE=95004649; PubMed=7920699;  
 RA Klonus D., Hoefgen R., Willmitzer L., Riesmeier J.W.;  
 RT "Isolation and characterization of two cDNAs encoding ATP-  
 sulfurylase from potato by complementation of a yeast mutant.";  
 RL Plant J. 6:105-112(1994).  
 DR EMBL; X79053; CAA55655.1; -;  
 DR InterPro; IPR002650; ATP-sulfurylase.

Search completed: May 30, 2003, 12:32:18  
Job time : 89 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:08:55 ; Search time 73 Seconds  
(without alignments)  
1122.591 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSGIKKQKTKNQKSTNVY.....MAPKAWKLVTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID          | Description        |
|------------|--------|-------------|--------|-------------|--------------------|
| 1          | 3216   | 98.3        | 619    | 21 AAY79214 | Human transferase  |
| 2          | 3162.5 | 96.7        | 628    | 22 AAU32438 | Novel human secret |
| 3          | 2602   | 79.5        | 624    | 19 AAW70494 | Human disease rela |
| 4          | 2602   | 79.5        | 624    | 20 AAY22349 | Human APS kinase/A |
| 5          | 2602   | 79.5        | 625    | 20 AAW67882 | Human secreted pro |
| 6          | 2031.5 | 62.1        | 635    | 22 ABB71650 | Drosophila melanog |
| 7          | 1294   | 39.6        | 460    | 21 AAG44040 | Arabidopsis thalia |
| 8          | 1294   | 39.6        | 463    | 21 AAG44039 | Arabidopsis thalia |
| 9          | 1282   | 39.2        | 462    | 21 AAG30254 | Arabidopsis thalia |
| 10         | 1282   | 39.2        | 465    | 21 AAG30253 | Arabidopsis thalia |

|    |        |      |     |             |                     |
|----|--------|------|-----|-------------|---------------------|
| 11 | 1282   | 39.2 | 510 | 21 AAG30252 | Arabidopsis thalia  |
| 12 | 1272.5 | 38.9 | 473 | 21 AAG09595 | Arabidopsis thalia  |
| 13 | 1272.5 | 38.9 | 476 | 21 AAG09594 | Arabidopsis thalia  |
| 14 | 1270   | 38.8 | 425 | 21 AAG09596 | Arabidopsis thalia  |
| 15 | 1259.5 | 38.5 | 408 | 21 AAG43667 | Arabidopsis thalia  |
| 16 | 1259.5 | 38.5 | 469 | 21 AAG43666 | Arabidopsis thalia  |
| 17 | 1252.5 | 38.3 | 396 | 21 AAG43668 | Arabidopsis thalia  |
| 18 | 1228   | 37.5 | 373 | 21 AAG44041 | Arabidopsis thalia  |
| 19 | 1225.5 | 37.5 | 446 | 21 AAG21121 | Arabidopsis thalia  |
| 20 | 1223.5 | 37.4 | 385 | 21 AAG21122 | Arabidopsis thalia  |
| 21 | 1216.5 | 37.2 | 373 | 21 AAG21123 | Arabidopsis thalia  |
| 22 | 616    | 18.8 | 174 | 21 AAG20074 | Arabidopsis thalia  |
| 23 | 581    | 17.8 | 166 | 21 AAG20075 | Arabidopsis thalia  |
| 24 | 508.5  | 15.5 | 311 | 21 AAY44791 | Soybean Adenylsulph |
| 25 | 499.5  | 15.3 | 246 | 21 AAY44792 | Wheat Adenylsulph   |
| 26 | 497.5  | 15.2 | 293 | 21 AAG29425 | Arabidopsis thalia  |
| 27 | 497.5  | 15.2 | 293 | 21 AAY77957 | A. thaliana enviro  |
| 28 | 482    | 14.7 | 224 | 21 AAY44788 | Corn Adenylsulph    |
| 29 | 482    | 14.7 | 343 | 21 AAY44789 | Corn Adenylsulph    |
| 30 | 481    | 14.7 | 204 | 23 ABP39672 | Staphylococcus epi  |
| 31 | 476.5  | 14.6 | 201 | 22 AAG98409 | Escherichia coli p  |
| 32 | 476    | 14.6 | 236 | 21 AAG14478 | Arabidopsis thalia  |
| 33 | 476    | 14.6 | 251 | 21 AAG11017 | Arabidopsis thalia  |
| 34 | 476    | 14.6 | 251 | 21 AAG38348 | Arabidopsis thalia  |
| 35 | 476    | 14.6 | 259 | 21 AAG14477 | Arabidopsis thalia  |
| 36 | 476    | 14.6 | 263 | 21 AAG11016 | Arabidopsis thalia  |
| 37 | 476    | 14.6 | 263 | 21 AAG38347 | Arabidopsis thalia  |
| 38 | 476    | 14.6 | 272 | 21 AAG14476 | Arabidopsis thalia  |
| 39 | 476    | 14.6 | 276 | 21 AAG11015 | Arabidopsis thalia  |
| 40 | 476    | 14.6 | 276 | 21 AAG38346 | Arabidopsis thalia  |
| 41 | 472.5  | 14.4 | 208 | 21 AAG35788 | Arabidopsis thalia  |
| 42 | 472    | 14.4 | 140 | 21 AAG20076 | Arabidopsis thalia  |
| 43 | 458.5  | 14.0 | 252 | 21 AAG14580 | Arabidopsis thalia  |
| 44 | 458.5  | 14.0 | 252 | 21 AAG47458 | Arabidopsis thalia  |
| 45 | 458.5  | 14.0 | 305 | 21 AAG14579 | Arabidopsis thalia  |

ALIGNMENTS

RESULT 1

AAV79214

ID AAY79214 standard; Protein; 619 AA.

XX AC AAY79214;

XX DT 19-JUN-2000 (first entry)

XX DE Human transferase TRNSFS-6.

XX KW Transferase; TRNSFS-6; human; antitumour; cancer;

KW gastrointestinal disorder; developmental disorder;

KW genetic disorder; neurological disorder; reproductive disorder;

KW smooth muscle disorder; immunological disorder; inflammation;

XX KW diagnosis; therapy; ATP sulfurylase/APS kinase 2.

XX OS Homo sapiens.

XX FH Key

XX FT Modified-site 92

XX FT Modified-site 176

XX FT Modified-site 180

XX FT Modified-site 198

XX FT Modified-site 250

XX FT Modified-site 285

XX FT Modified-site 308

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

XX FT Modified-site /note= "potential O-phosphorylation"

|    |               |        |  |          |   |   |     |
|----|---------------|--------|--|----------|---|---|-----|
| FT | Modified-site | 313    | /note= "potential O-phosphorylation"   | QY       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 394    | /note= "potential O-phosphorylation"   | Db       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 453    | /note= "potential O-phosphorylation"   | QY       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 527    | /note= "potential O-phosphorylation"   | Db       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 548    | /note= "potential O-phosphorylation"   | QY       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDD-----DGVINMS   | 295 |
| FT | Modified-site | 430    | /note= "potential O-phosphorylation"   | Db       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDDMALPDGVINMS    | 300 |
| FT | Modified-site | 439    | /note= "potential O-phosphorylation"   | QY       | 296   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| FT | Modified-site | 195    | /note= "potential N-glycosylation"     | Db       | 301   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| FT | Modified-site | 298    | /note= "potential N-glycosylation"     | QY       | 356   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 415 |
| FT | Binding-site  | 49..56 | /note= "ATP/GTP binding site (P-loop)" | Db       | 361   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 420 |
| XX |               |        |  | QY       | 416   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 475 |
| XX |               |        |  | Db       | 421   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 480 |
| XX |               |        |  | QY       | 476   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 535 |
| XX |               |        |  | Db       | 481   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 540 |
| XX |               |        |  | QY       | 536   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 595 |
| XX |               |        |  | Db       | 541   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 600 |
| XX |               |        |  | QY       | 596   | MAPKAWKVLTDYRSEMDXN 615                                     |     |
| XX |               |        |  | Db       | 601   | MAPKAWKVLTDYRS-LEKN 619                                     |     |
| XX |               |        |  | RESULT 2 |   |   |     |
| XX |               |        |  | AAU32438 |   |   |     |
| XX |               |        |  | ID       | AAU32438  | standard; Protein; 648 AA.                                  |     |
| XX |               |        |  | AC       | AAU32438;   |   |     |
| XX |               |        |  | DT       | 18-DEC-2001   | (first entry)   |     |
| XX |               |        |  | DE       | Novel human secreted protein #2929.                                     |   |     |
| XX |               |        |  | KW       | Human: vaccination; gene therapy; nutritional supplement;               |   |     |
| XX |               |        |  | KW       | stem cell proliferation; haematopoiesis; nerve tissue regeneration;     |   |     |
| XX |               |        |  | KW       | immune suppression; immune stimulation; anti-inflammatory; leukaemia.   |   |     |
| XX |               |        |  | OS       | Homo sapiens.   |   |     |
| XX |               |        |  | PN       | WO200179449-A2.   |   |     |
| XX |               |        |  | PD       | 25-OCT-2001.  |   |     |
| XX |               |        |  | PF       | 16-APR-2001; 2001WO-US08656.  |   |     |
| XX |               |        |  | PR       | 18-APR-2000; 2000US-0552929.  |   |     |
| XX |               |        |  | PR       | 26-JAN-2001; 2001US-0770160.  |   |     |
| XX |               |        |  | PA       | (HYSE-) HYSEQ INC.  |   |     |
| XX |               |        |  | PI       | Tang YT, Liu C, Drmanac RT;   |   |     |
| XX |               |        |  | DR       | WPI; 2001-611725/70.  |   |     |
| XX |               |        |  | PT       | Nucleic acids encoding a range of human polypeptides, useful in genetic |   |     |
| XX |               |        |  | PT       | vaccination, testing and therapy -                                      |   |     |
| XX |               |        |  | XX       | Claim 20; Page 613; 765pp; English.                                     |   |     |

|    |               |        |  |          |   |   |     |
|----|---------------|--------|--|----------|---|---|-----|
| FT | Modified-site | 313    | /note= "potential O-phosphorylation"   | QY       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 394    | /note= "potential O-phosphorylation"   | Db       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 453    | /note= "potential O-phosphorylation"   | QY       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 527    | /note= "potential O-phosphorylation"   | Db       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 548    | /note= "potential O-phosphorylation"   | QY       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDD-----DGVINMS   | 295 |
| FT | Modified-site | 430    | /note= "potential O-phosphorylation"   | Db       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDDMALPDGVINMS    | 300 |
| FT | Modified-site | 439    | /note= "potential O-phosphorylation"   | QY       | 296   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| FT | Modified-site | 195    | /note= "potential N-glycosylation"     | Db       | 301   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| FT | Modified-site | 298    | /note= "potential N-glycosylation"     | QY       | 356   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 415 |
| FT | Binding-site  | 49..56 | /note= "ATP/GTP binding site (P-loop)" | Db       | 361   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 420 |
| XX |               |        |  | QY       | 416   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 475 |
| XX |               |        |  | Db       | 421   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 480 |
| XX |               |        |  | QY       | 476   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 535 |
| XX |               |        |  | Db       | 481   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 540 |
| XX |               |        |  | QY       | 536   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 595 |
| XX |               |        |  | Db       | 541   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 600 |
| XX |               |        |  | QY       | 596   | MAPKAWKVLTDYRSEMDXN 615                                     |     |
| XX |               |        |  | Db       | 601   | MAPKAWKVLTDYRS-LEKN 619                                     |     |
| XX |               |        |  | RESULT 2 |   |   |     |
| XX |               |        |  | AAU32438 |   |   |     |
| XX |               |        |  | ID       | AAU32438  | standard; Protein; 648 AA.                                  |     |
| XX |               |        |  | AC       | AAU32438;   |   |     |
| XX |               |        |  | DT       | 18-DEC-2001   | (first entry)   |     |
| XX |               |        |  | DE       | Novel human secreted protein #2929.                                     |   |     |
| XX |               |        |  | KW       | Human: vaccination; gene therapy; nutritional supplement;               |   |     |
| XX |               |        |  | KW       | stem cell proliferation; haematopoiesis; nerve tissue regeneration;     |   |     |
| XX |               |        |  | KW       | immune suppression; immune stimulation; anti-inflammatory; leukaemia.   |   |     |
| XX |               |        |  | OS       | Homo sapiens.   |   |     |
| XX |               |        |  | PN       | WO200179449-A2.   |   |     |
| XX |               |        |  | PD       | 25-OCT-2001.  |   |     |
| XX |               |        |  | PF       | 16-APR-2001; 2001WO-US08656.  |   |     |
| XX |               |        |  | PR       | 18-APR-2000; 2000US-0552929.  |   |     |
| XX |               |        |  | PR       | 26-JAN-2001; 2001US-0770160.  |   |     |
| XX |               |        |  | PA       | (HYSE-) HYSEQ INC.  |   |     |
| XX |               |        |  | PI       | Tang YT, Liu C, Drmanac RT;   |   |     |
| XX |               |        |  | DR       | WPI; 2001-611725/70.  |   |     |
| XX |               |        |  | PT       | Nucleic acids encoding a range of human polypeptides, useful in genetic |   |     |
| XX |               |        |  | PT       | vaccination, testing and therapy -                                      |   |     |
| XX |               |        |  | XX       | Claim 20; Page 613; 765pp; English.                                     |   |     |

|    |               |        |  |          |   |   |     |
|----|---------------|--------|--|----------|---|---|-----|
| FT | Modified-site | 313    | /note= "potential O-phosphorylation"   | QY       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 394    | /note= "potential O-phosphorylation"   | Db       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 453    | /note= "potential O-phosphorylation"   | QY       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 527    | /note= "potential O-phosphorylation"   | Db       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 548    | /note= "potential O-phosphorylation"   | QY       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDD-----DGVINMS   | 295 |
| FT | Modified-site | 430    | /note= "potential O-phosphorylation"   | Db       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDDMALPDGVINMS    | 300 |
| FT | Modified-site | 439    | /note= "potential O-phosphorylation"   | QY       | 296   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| FT | Modified-site | 195    | /note= "potential N-glycosylation"     | Db       | 301   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| FT | Modified-site | 298    | /note= "potential N-glycosylation"     | QY       | 356   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 415 |
| FT | Binding-site  | 49..56 | /note= "ATP/GTP binding site (P-loop)" | Db       | 361   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 420 |
| XX |               |        |  | QY       | 416   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 475 |
| XX |               |        |  | Db       | 421   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 480 |
| XX |               |        |  | QY       | 476   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 535 |
| XX |               |        |  | Db       | 481   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 540 |
| XX |               |        |  | QY       | 536   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 595 |
| XX |               |        |  | Db       | 541   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 600 |
| XX |               |        |  | QY       | 596   | MAPKAWKVLTDYRSEMDXN 615                                     |     |
| XX |               |        |  | Db       | 601   | MAPKAWKVLTDYRS-LEKN 619                                     |     |
| XX |               |        |  | RESULT 2 |   |   |     |
| XX |               |        |  | AAU32438 |   |   |     |
| XX |               |        |  | ID       | AAU32438  | standard; Protein; 648 AA.                                  |     |
| XX |               |        |  | AC       | AAU32438;   |   |     |
| XX |               |        |  | DT       | 18-DEC-2001   | (first entry)   |     |
| XX |               |        |  | DE       | Novel human secreted protein #2929.                                     |   |     |
| XX |               |        |  | KW       | Human: vaccination; gene therapy; nutritional supplement;               |   |     |
| XX |               |        |  | KW       | stem cell proliferation; haematopoiesis; nerve tissue regeneration;     |   |     |
| XX |               |        |  | KW       | immune suppression; immune stimulation; anti-inflammatory; leukaemia.   |   |     |
| XX |               |        |  | OS       | Homo sapiens.   |   |     |
| XX |               |        |  | PN       | WO200179449-A2.   |   |     |
| XX |               |        |  | PD       | 25-OCT-2001.  |   |     |
| XX |               |        |  | PF       | 16-APR-2001; 2001WO-US08656.  |   |     |
| XX |               |        |  | PR       | 18-APR-2000; 2000US-0552929.  |   |     |
| XX |               |        |  | PR       | 26-JAN-2001; 2001US-0770160.  |   |     |
| XX |               |        |  | PA       | (HYSE-) HYSEQ INC.  |   |     |
| XX |               |        |  | PI       | Tang YT, Liu C, Drmanac RT;   |   |     |
| XX |               |        |  | DR       | WPI; 2001-611725/70.  |   |     |
| XX |               |        |  | PT       | Nucleic acids encoding a range of human polypeptides, useful in genetic |   |     |
| XX |               |        |  | PT       | vaccination, testing and therapy -                                      |   |     |
| XX |               |        |  | XX       | Claim 20; Page 613; 765pp; English.                                     |   |     |

|    |               |        |  |          |   |   |     |
|----|---------------|--------|--|----------|---|---|-----|
| FT | Modified-site | 313    | /note= "potential O-phosphorylation"   | QY       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 394    | /note= "potential O-phosphorylation"   | Db       | 121   | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 453    | /note= "potential O-phosphorylation"   | QY       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 527    | /note= "potential O-phosphorylation"   | Db       | 181   | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 548    | /note= "potential O-phosphorylation"   | QY       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDD-----DGVINMS   | 295 |
| FT | Modified-site | 430    | /note= "potential O-phosphorylation"   | Db       | 241   | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDDMALPDGVINMS    | 300 |
| FT | Modified-site | 439    | /note= "potential O-phosphorylation"   | QY       | 296   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| FT | Modified-site | 195    | /note= "potential N-glycosylation"     | Db       | 301   | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| FT | Modified-site | 298    | /note= "potential N-glycosylation"     | QY       | 356   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 415 |
| FT | Binding-site  | 49..56 | /note= "ATP/GTP binding site (P-loop)" | Db       | 361   | IKVMESGDMVGGDLQVLEKIRWNGDLDQYRLTPELKKCKCKEMNADAVFAFOLRNPVH  | 420 |
| XX |               |        |  | QY       | 416   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 475 |
| XX |               |        |  | Db       | 421   | NGHALIMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKS  | 480 |
| XX |               |        |  | QY       | 476   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 535 |
| XX |               |        |  | Db       | 481   | TIVAIPPSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMPHETKDLIYPTGGKV   | 540 |
| XX |               |        |  | QY       | 536   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 595 |
| XX |               |        |  | Db       | 541   | LSMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDIFISGTRMKLAREGENPDGF     | 600 |
| XX |               |        |  | QY       | 596   | MAPKAWKVLTDYRSEMDXN 615                                     |     |
| XX |               |        |  | Db       | 601   | MAPKAWKVLTDYRS-LEKN 619                                     |     |
| XX |               |        |  | RESULT 2 |   |   |     |
| XX |               |        |  | AAU32438 |   |   |     |
| XX |               |        |  | ID       | AAU32438  | standard; Protein; 648 AA.                                  |     |
| XX |               |        |  | AC       | AAU32438;   |   |     |
| XX |               |        |  | DT       | 18-DEC-2001   | (first entry)   |     |
| XX |               |        |  | DE       | Novel human secreted protein #2929.                                     |   |     |
| XX |               |        |  | KW       | Human: vaccination; gene therapy; nutritional supplement;               |   |     |
| XX |               |        |  | KW       | stem cell proliferation; haematopoiesis; nerve tissue regeneration;     |   |     |
| XX |               |        |  | KW       | immune suppression; immune stimulation; anti-inflammatory; leukaemia.   |   |     |
| XX |               |        |  | OS       | Homo sapiens.   |   |     |
| XX |               |        |  | PN       | WO200179449-A2.   |   |     |
| XX |               |        |  | PD       | 25-OCT-2001.  |   |     |
| XX |               |        |  | PF       | 16-APR-2001; 2001WO-US08656.  |   |     |
| XX |               |        |  | PR       | 18-APR-2000; 2000US-0552929.  |   |     |
| XX |               |        |  | PR       | 26-JAN-2001; 2001US-0770160.  |   |     |
| XX |               |        |  | PA       | (HYSE-) HYSEQ INC.  |   |     |
| XX |               |        |  | PI       | Tang YT, Liu C, Drmanac RT;   |   |     |
| XX |               |        |  | DR       | WPI; 2001-611725/70.  |   |     |
| XX |               |        |  | PT       | Nucleic acids encoding a range of human polypeptides, useful in genetic |   |     |
| XX |               |        |  | PT       | vaccination, testing and therapy -                                      |   |     |
| XX |               |        |  | XX       | Claim 20; Page 613; 765pp; English.                                     |   |     |

|    |               |         |                                       |    |     |   |     |
|----|---------------|---------|---------------------------------------|----|-----|---|-----|
| FT | Modified-site | 313     | /note="potential O-phosphorylation"   | QY | 1   | MSGIKKKTENQOKSTNVVYQAHVSRNKRQGVVGRGFRGCTVWL7GLSGAGKTTISF    | 60  |
| FT | Modified-site | 394     | /note="potential O-phosphorylation"   | Db | 1   | MSGIKKKTENQOKSTNVVYQAHVSRNKRQGVVGRGFRGCTVWL7GLSGAGKTTISF    | 60  |
| FT | Modified-site | 453     | /note="potential O-phosphorylation"   | QY | 61  | ALEEVLVSHAIPCYSLDGDVNRGLNRLNLFSPGDREENIRIAEVAKLFDAGLVCLITS  | 120 |
| FT | Modified-site | 527     | /note="potential O-phosphorylation"   | Db | 61  | ALEEVLVSHAIPCYSLDGDVNRGLNRLNLFSPGDREENIRIAEVAKLFDAGLVCLITS  | 120 |
| FT | Modified-site | 548     | /note="potential O-phosphorylation"   | QY | 121 | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 430     | /note="potential O-phosphorylation"   | Db | 121 | FISPPAKDRENARKIHESAGLFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS | 180 |
| FT | Modified-site | 439     | /note="potential O-phosphorylation"   | QY | 181 | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 195     | /note="potential N-glycosylation"     | Db | 181 | DYEKPEPTELKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDIHELFPENKLDHVAE   | 240 |
| FT | Modified-site | 298     | /note="potential N-glycosylation"     | QY | 241 | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDD-----DGVINMS   | 295 |
| FT | Binding-site  | 49...56 | /note="ATP/GTP binding site (P-loop)" | Db | 241 | AETPLSLITKLDLQWQVLSGCVATPLKGFMEKEYLQVMHFDTLDDMALPDGVINMS    | 300 |
| FT |               |         |                                       | QY | 296 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 296 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       | QY | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 355 |
| XX |               |         |                                       | Db | 356 | IPVLPVSAEDKTRLEGCSKFVLAHGRRVAILRDAEFYEHKREKCSRVMGTCTCKHPH   | 360 |
| XX |               |         |                                       |    |     |   |     |



Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTI 58  
 DB 9 KKVLSNNAQNWGMQRATNTYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTV 68

QY 59 SFALEEYLVSHAIPCYSLDGDNVHGLNRNLGSPGDRREENIRIARIVAKLFDAGLVCI 118  
 DB 69 SMALEEYLVCHGIPCYTLTDGDNIRQGLNKNLGFSPEDREENVRRIAEVAKLFDAGLVCI 128

QY 119 TSFISPAKRENNARKIHESAGLPFFEIFVDAPLNICESDVGLYKRRARAGEIKGFTGI 178  
 DB 129 TSFISPTQDRNNARQIHESAGLPFFEIFVDAPLHCEQDRVGLYKRRARAGEIKGFTGI 188

QY 179 DSYEKPTPERVLKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDHELEVPENKLDHVR 238  
 DB 189 DSEYKPEAPELVKTDSCDVNDVCQVQVVELLQEQNIIVPYTIIKDHELEVPENKLDHVR 248

QY 239 AEATLPSLITKLDLQWQVLSGWTPLKGFMRKEYLOVMHFDTLDDGVNINSIPI 298  
 DB 249 TDAETLPALKINKVDQWQVLAEGWATPLNGFMREYQLCLHFDCLDGGVINSVPI 308

QY 299 VLPVSAEDKRLKLECGSFVLAHGGRRVAILRDAEFYEHKREKRSRVWGTCTCKHPHIKM 358  
 DB 309 VLTATHEDKRLDGCCTAFALMYEGRRVAILRNPFEFFEHKREKRCARQWGTCKNHPYIKM 368

QY 359 VMESGDLVGGDLQVLEKIRWNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 418  
 DB 369 VMEQGDWLIGDLQVLDVRVYNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRLLERGYKHPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 478  
 DB 429 ALLMODTHKOLLERGYRRPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPHGGKVLMS 538  
 DB 489 AIFPSPMYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPHGGKVLMS 548

QY 539 APLGTSVEIIPFRVAAYNKKAKKAMDFYDPAHNEFDIFISGTRMKRLAREGENPPDGFMAP 598  
 DB 549 APLGLTLEIVPFRVAAYNKKKRMDDYDSEHHEDEFFISGTRMKRLAREQKPPDGFMAP 608

QY 599 KAWKVLTDYKRS 610  
 DB 609 KAWTVLLEYKRS 620

RESULT 4  
 AAY22349  
 ID AAY22349 standard; Protein; 624 AA.

XX AAY22349;

XX 24-SEP-1999 (first entry)

DE Human APS kinase/ATP sulphurylase protein.

XX APS kinase/ATP sulphurylase; human; PAPS production;  
 KW 3'-phosphoadenosine 5'-phosphosulphate.

OS Homo sapiens.

XX JP11187883-A.

XX 13-JUL-1999.

XX 26-DEC-1997; 97JP-0360387.

XX 26-DEC-1997; 97JP-0360387.

XX (NIRA ) UNITIKA LTD.

PA (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.

XX

DR WPI; 1999-451549/38.  
 DR N-PSDB; AAX84897.  
 XX New human-derived APS kinase/ATP sulphurylase gene - useful for  
 PT large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate  
 PT (PAPS)  
 XX Claim 1; Page 6-7; 9pp; Japanese.  
 PS This sequence is the human-derived APS kinase/ATP sulphurylase of  
 CC the invention. The enzyme may be used to prepare 3'-phosphoadenosine  
 CC 5'-phosphosulphate (PAPS) on a large scale.  
 XX

SQ Sequence 624 AA;

Query Match 79.5%; Score 2602; DB 20; Length 624;  
 Best Local Similarity 77.1%; Pred. No. 4.2e-254;  
 Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTI 58  
 DB 9 KKVLSNNAQNWGMQRATNTYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTV 68

QY 59 SFALEEYLVSHAIPCYSLDGDNVHGLNRNLGSPGDRREENIRIARIVAKLFDAGLVCI 118  
 DB 69 SMALEEYLVCHGIPCYTLTDGDNIRQGLNKNLGFSPEDREENVRRIAEVAKLFDAGLVCI 128

QY 119 TSFISPAKRENNARKIHESAGLPFFEIFVDAPLNICESDVGLYKRRARAGEIKGFTGI 178  
 DB 129 TSFISPTQDRNNARQIHESAGLPFFEIFVDAPLHCEQDRVGLYKRRARAGEIKGFTGI 188

QY 179 DSYEKPTPERVLKTNLSTVSDCVHQVVELLQEQNIIVPYTIIKDHELEVPENKLDHVR 238  
 DB 189 DSEYKPEAPELVKTDSCDVNDVCQVQVVELLQEQNIIVPYTIIKDHELEVPENKLDHVR 248

QY 239 AEATLPSLITKLDLQWQVLSGWTPLKGFMRKEYLOVMHFDTLDDGVNINSIPI 298  
 DB 249 TDAETLPALKINKVDQWQVLAEGWATPLNGFMREYQLCLHFDCLDGGVINSVPI 308

QY 299 VLPVSAEDKRLKLECGSFVLAHGGRRVAILRDAEFYEHKREKRSRVWGTCTCKHPHIKM 358  
 DB 309 VLTATHEDKRLDGCCTAFALMYEGRRVAILRNPFEFFEHKREKRCARQWGTCKNHPYIKM 368

QY 359 VMESGDLVGGDLQVLEKIRWNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 418  
 DB 369 VMEQGDWLIGDLQVLDVRVYNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRLLERGYKHPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 478  
 DB 429 ALLMODTHKOLLERGYRRPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPHGGKVLMS 538  
 DB 489 AIFPSPMYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPHGGKVLMS 548

QY 539 APLGTSVEIIPFRVAAYNKKAKKAMDFYDPAHNEFDIFISGTRMKRLAREGENPPDGFMAP 598  
 DB 549 APLGLTLEIVPFRVAAYNKKKRMDDYDSEHHEDEFFISGTRMKRLAREQKPPDGFMAP 608

QY 599 KAWKVLTDYKRS 610  
 DB 609 KAWTVLLEYKRS 620

RESULT 5  
 AAW67882

ID AAW67882 standard; Protein; 625 AA.

XX AAW67882;

XX 25-MAR-1999 (first entry)

XX Human secreted protein encoded by gene 76 clone HOSFD58.





PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139730.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147036.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149729.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 39.6%; Score 1294; DB 21; Length 460;

Best Local Similarity 62.2%; Pred. No. 8.8e-122;

Matches 242; Conservative 61; Mismatches 82; Indels 4; Gaps 4;

QY 226 ELFPENKLDHVRRAEAEETPLSLITKLDQWQVLSGWTATPLKGFMRKEYLQVMHFDT 285

DB 57 ELIVEEPKRREKKHEAADLPRVELTAIDLQWHLVLSGWTATPLKGFMRSEFLQTLHNS 116

QY 286 L-LDDG-VINMSIPIVLPVSAEDKTRLECCSKFVLAHG-GRRVAILRDAEYEHKKEERC 342

|          |  |  |     |    |              |               |
|----------|--|--|-----|----|--------------|---------------|
| Db       | 117  | LRLLDGSVWMSPIVLADDEQKARIGESTRVALENSDGNPVSILSDIYIYKHPKEERI    | 176 | PR | 18-MAY-1999; | 99US-0134768. |
| Qy       | 343  | SRVWGTCTTKHPH-KVMESGDLVGGDLQVLEKIRWNDGLDQYRLTPLELKQCKEMN     | 401 | PR | 19-MAY-1999; | 99US-0134941. |
| Db       | 177  | ARTWGTAPGLPVDEAITWAGNWLIGGDLVLEPVKYINDGLDRFLSPALRKLEKRN      | 236 | PR | 20-MAY-1999; | 99US-0135124. |
| Qy       | 402  | ADAVEAFQLRNPVHNGHALLMDTCRLRLRGYKHPVLLHPLGGWTKDDVPLDWRMKQ     | 461 | PR | 21-MAY-1999; | 99US-0135353. |
| Db       | 237  | ADAVEAFQLRNPVHNGHALLMDTCRLRLRGYKHPVLLHPLGGWTKDDVPLDWRMKQ     | 296 | PR | 24-MAY-1999; | 99US-0135829. |
| Qy       | 462  | HAAVLEGVLDPKSTIVAPPSMLYAGTPEVOWHCRSRMIAGANFYIVGRDPAGMPHPE    | 521 | PR | 25-MAY-1999; | 99US-0136021. |
| Db       | 297  | HEKVLDEGVLDPEPTTVYSIPSPMHYAGTPEVOWHAKARINAGANFYIVGRDPAGMPHVP | 356 | PR | 27-MAY-1999; | 99US-0136392. |
| Qy       | 522  | TKKDLVEPTHGGKVLNMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNERDFISGTRM     | 591 | PR | 01-JUN-1999; | 99US-0136782. |
| Db       | 357  | EKRDLDVADHGKVLNMAPGLERINILPFRVAAYDKTOGKMAFFDFSRQDFLFISGTRM   | 416 | PR | 03-JUN-1999; | 99US-0137528. |
| Qy       | 582  | RKLAREGENPPDGFMAPKAMKVLTDYYS 610                             |     | PR | 04-JUN-1999; | 99US-0137502. |
| Db       | 417  | RTLAKNNENPPDGFMCPPGGKVLVDYVES 445                            |     | PR | 07-JUN-1999; | 99US-0137724. |
| RESULT 8 |  |  |     | PR | 08-JUN-1999; | 99US-0138094. |
| AAG44039 |  |  |     | PR | 10-JUN-1999; | 99US-0138540. |
| ID       | AAG44039   | standard; Protein: 463 AA.                                   |     | PR | 10-JUN-1999; | 99US-0138847. |
| XX       |  |  |     | PR | 14-JUN-1999; | 99US-0139119. |
| AC       | AAG44039;  |  |     | PR | 16-JUN-1999; | 99US-0139452. |
| DT       | 18-OCT-2000  | (first entry)  |     | PR | 16-JUN-1999; | 99US-0139453. |
| XX       |  |  |     | PR | 17-JUN-1999; | 99US-0139492. |
| DE       | Arabidopsis thaliana protein fragment SEQ ID NO: 55115.                  |  |     | PR | 18-JUN-1999; | 99US-0139454. |
| KW       | Protein identification; signal transduction pathway; metabolic pathway;  |  |     | PR | 18-JUN-1999; | 99US-0139455. |
| KW       | hybridisation assay; genetic mapping; gene expression control; promoter; |  |     | PR | 18-JUN-1999; | 99US-0139456. |
| KW       | termination sequence.  |  |     | PR | 18-JUN-1999; | 99US-0139461. |
| XX       |  |  |     | PR | 18-JUN-1999; | 99US-0139462. |
| OS       | Arabidopsis thaliana.  |  |     | PR | 18-JUN-1999; | 99US-0139463. |
| PN       | EP1033405-A2.  |  |     | PR | 18-JUN-1999; | 99US-0139464. |
| XX       |  |  |     | PR | 18-JUN-1999; | 99US-0139465. |
| XX       | 06-SEP-2000.   |  |     | PR | 18-JUN-1999; | 99US-0139466. |
| XX       | 25-FEB-2000; 2000EP-0301439.   |  |     | PR | 18-JUN-1999; | 99US-0139467. |
| XX       |  |  |     | PR | 18-JUN-1999; | 99US-0139468. |
| PR       | 25-FEB-1999;   | 99US-0121825.  |     | PR | 18-JUN-1999; | 99US-0139469. |
| PR       | 05-MAR-1999;   | 99US-0123180.  |     | PR | 18-JUN-1999; | 99US-0139470. |
| PR       | 03-MAR-1999;   | 99US-0123548.  |     | PR | 18-JUN-1999; | 99US-0139471. |
| PR       | 23-MAR-1999;   | 99US-0125788.  |     | PR | 18-JUN-1999; | 99US-0139472. |
| PR       | 25-MAR-1999;   | 99US-0126264.  |     | PR | 18-JUN-1999; | 99US-0139473. |
| PR       | 29-MAR-1999;   | 99US-0126785.  |     | PR | 18-JUN-1999; | 99US-0139474. |
| PR       | 01-APR-1999;   | 99US-0127462.  |     | PR | 18-JUN-1999; | 99US-0139475. |
| PR       | 06-APR-1999;   | 99US-0128234.  |     | PR | 18-JUN-1999; | 99US-0139476. |
| PR       | 08-APR-1999;   | 99US-0128714.  |     | PR | 18-JUN-1999; | 99US-0139477. |
| PR       | 16-APR-1999;   | 99US-0129845.  |     | PR | 18-JUN-1999; | 99US-0139478. |
| PR       | 19-APR-1999;   | 99US-0130077.  |     | PR | 18-JUN-1999; | 99US-0139479. |
| PR       | 21-APR-1999;   | 99US-0130449.  |     | PR | 18-JUN-1999; | 99US-0139480. |
| PR       | 23-APR-1999;   | 99US-0130510.  |     | PR | 18-JUN-1999; | 99US-0139481. |
| PR       | 28-APR-1999;   | 99US-0130891.  |     | PR | 18-JUN-1999; | 99US-0139482. |
| PR       | 30-APR-1999;   | 99US-0131449.  |     | PR | 18-JUN-1999; | 99US-0139483. |
| PR       | 30-APR-1999;   | 99US-0132048.  |     | PR | 18-JUN-1999; | 99US-0139484. |
| PR       | 04-MAY-1999;   | 99US-0132407.  |     | PR | 18-JUN-1999; | 99US-0139485. |
| PR       | 05-MAY-1999;   | 99US-0132484.  |     | PR | 18-JUN-1999; | 99US-0139486. |
| PR       | 06-MAY-1999;   | 99US-0132485.  |     | PR | 18-JUN-1999; | 99US-0139487. |
| PR       | 07-MAY-1999;   | 99US-0132486.  |     | PR | 18-JUN-1999; | 99US-0139488. |
| PR       | 11-MAY-1999;   | 99US-0132863.  |     | PR | 18-JUN-1999; | 99US-0139489. |
| PR       | 14-MAY-1999;   | 99US-0134256.  |     | PR | 18-JUN-1999; | 99US-0139490. |
| PR       | 14-MAY-1999;   | 99US-0134218.  |     | PR | 18-JUN-1999; | 99US-0139491. |
| PR       | 14-MAY-1999;   | 99US-0134219.  |     | PR | 18-JUN-1999; | 99US-0139492. |
| PR       | 14-MAY-1999;   | 99US-0134221.  |     | PR | 18-JUN-1999; | 99US-0139493. |
| PR       | 14-MAY-1999;   | 99US-0134370.  |     | PR | 18-JUN-1999; | 99US-0139494. |



| PR                        | 26-OCT-1999;                 | 99US-0161359;        |
|---------------------------|------------------------------|----------------------|
| PR                        | 26-OCT-1999;                 | 99US-0161360.        |
| PR                        | 26-OCT-1999;                 | 99US-0161361.        |
| PR                        | 28-OCT-1999;                 | 99US-0161920.        |
| PR                        | 28-OCT-1999;                 | 99US-0161920.        |
| PR                        | 28-OCT-1999;                 | 99US-0161993.        |
| PR                        | 28-OCT-1999;                 | 99US-0162142.        |
| PR                        | 29-OCT-1999;                 | 99US-0162142.        |
| Query Match               |                              |                      |
| Best Local Similarity     |                              | 39.68;               |
| Matches 242; Conservative |                              | 62.28;               |
| Matches 242; Conservative |                              |                      |
| QY                        | 226                          | ELFVPEKNKLDHVRAEATLL |
| DB                        | 60                           | ELIVEEPKREKKRHEADLL  |
| QY                        | 286                          | L-LDDG-VINMSIPVLPLV  |
| DB                        | 120                          | LRDDGSGVNSVPILALAI   |
| QY                        | 343                          | SRVNGTCTCKRPHI-KVMMI |
| DB                        | 180                          | ARTWGTAPGLPYVDEAIT   |
| QY                        | 402                          | ADAVFAQLRNPVNHGHAL   |
| DB                        | 240                          | ADAVFAQLRNPVNHGHAL   |
| QY                        | 462                          | HAAVLEBEGVLDPKSTIVAI |
| DB                        | 300                          | HEKVLDDGVLDPEPTVVSII |
| QY                        | 522                          | TKKDLVETHTGGKVLSPAP  |
| DB                        | 360                          | EKRDLVDADHGKVKVLSPAP |
| QY                        | 582                          | RKLAREGNPPDQPMAPKPA  |
| DB                        | 420                          | RTLAKNNENPPDQPMCPGG  |
| RESULT 9                  |                              |                      |
| AAG30254                  |                              |                      |
| ID                        | AAG30254 standard; Protein:  |                      |
| XX                        | AAG30254;                    |                      |
| XX                        | 17-OCT-2000 (first entry)    |                      |
| DT                        | Arabidopsis thaliana protein |                      |
| DE                        | Arabidopsis thaliana.        |                      |
| XX                        | Protein identification; sign |                      |
| KW                        | hybridisation assay; genetic |                      |
| KW                        | termination sequence.        |                      |
| XX                        | Arabidopsis thaliana.        |                      |
| XX                        | EP1033405-A2.                |                      |
| PN                        | 06-SEP-2000.                 |                      |
| XX                        | 25-FEB-2000; 2000EP-0301439. |                      |
| XX                        | 25-FEB-1999; 99US-0121825.   |                      |
| PR                        | 05-MAR-1999; 99US-0123180.   |                      |
| PR                        | 09-MAR-1999; 99US-0123548.   |                      |
| PR                        | 23-MAR-1999; 99US-0125788.   |                      |
| PR                        | 25-MAR-1999; 99US-0126264.   |                      |
| PR                        | 29-MAR-1999; 99US-0126785.   |                      |
| PR                        | 01-APR-1999; 99US-0127462.   |                      |
| PR                        | 06-APR-1999; 99US-0128234.   |                      |
| PR                        | 08-APR-1999; 99US-0128714.   |                      |
| PR                        | 16-APR-1999; 99US-0129845.   |                      |

| PR                        | 26-OCT-1999;                 | 99US-0161359;        |
|---------------------------|------------------------------|----------------------|
| PR                        | 26-OCT-1999;                 | 99US-0161360.        |
| PR                        | 26-OCT-1999;                 | 99US-0161361.        |
| PR                        | 28-OCT-1999;                 | 99US-0161920.        |
| PR                        | 28-OCT-1999;                 | 99US-0161920.        |
| PR                        | 28-OCT-1999;                 | 99US-0161993.        |
| PR                        | 28-OCT-1999;                 | 99US-0162142.        |
| PR                        | 29-OCT-1999;                 | 99US-0162142.        |
| Query Match               |                              |                      |
| Best Local Similarity     |                              | 39.68;               |
| Matches 242; Conservative |                              | 62.28;               |
| Matches 242; Conservative |                              |                      |
| QY                        | 226                          | ELFVPEKNKLDHVRAEAEFL |
| DB                        | 60                           | ELIVEEPKREKKRHEADLL  |
| QY                        | 286                          | L-LDDG-VINMSDIPVLPLV |
| DB                        | 120                          | LRDDGSGVNMSPVILAI    |
| QY                        | 343                          | SRVNGTCTCKRPHI-KVMVM |
| DB                        | 180                          | ARTWGTAPGLPYVDEAIT   |
| QY                        | 402                          | ADAVFAQLRNPVNHGHAL   |
| DB                        | 240                          | ADAVFAQLRNPVNHGHAL   |
| QY                        | 462                          | HAAVLEBEGVLDPKSTIVAI |
| DB                        | 300                          | HEKVLDDGVLDPEPTVVSII |
| QY                        | 522                          | TKKDLVETPTGGKKVLSMAP |
| DB                        | 360                          | EKRDLVDADHGKKVLSMAP  |
| QY                        | 582                          | RKLAREGENPPDGPMAFKA  |
| DB                        | 420                          | RTLAKNNENPPDGPMPGG   |
| RESULT 9                  |                              |                      |
| AAG30254                  |                              |                      |
| ID                        | AAG30254 standard; Protein:  |                      |
| XX                        | AAG30254;                    |                      |
| XX                        | 17-OCT-2000 (first entry)    |                      |
| DT                        | Arabidopsis thaliana protein |                      |
| DE                        | Protein identification; sign |                      |
| XX                        | hybridisation assay; genetic |                      |
| KW                        | termination sequence.        |                      |
| XX                        | Arabidopsis thaliana.        |                      |
| XX                        | EP1033405-A2.                |                      |
| PN                        | 06-SEP-2000.                 |                      |
| XX                        | 25-FEB-2000; 2000EP-0301439. |                      |
| XX                        | 25-FEB-1999; 99US-0121825.   |                      |
| PR                        | 05-MAR-1999; 99US-0123180.   |                      |
| PR                        | 09-MAR-1999; 99US-0123548.   |                      |
| PR                        | 23-MAR-1999; 99US-0125788.   |                      |
| PR                        | 25-MAR-1999; 99US-0126264.   |                      |
| PR                        | 29-MAR-1999; 99US-0126785.   |                      |
| PR                        | 01-APR-1999; 99US-0127462.   |                      |
| PR                        | 06-APR-1999; 99US-0128234.   |                      |
| PR                        | 08-APR-1999; 99US-0128714.   |                      |
| PR                        | 16-APR-1999; 99US-0129845.   |                      |

PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-APR-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.



|    |              |               |    |              |               |
|----|--------------|---------------|----|--------------|---------------|
| PR | 01-JUL-1999; | 99US-0141842. | PR | 16-SEP-1999; | 99US-0154039. |
| PR | 01-JUL-1999; | 99US-0142154. | PR | 20-SEP-1999; | 99US-0154779. |
| PR | 02-JUL-1999; | 99US-0142055. | PR | 22-SEP-1999; | 99US-0155139. |
| PR | 06-JUL-1999; | 99US-0142390. | PR | 23-SEP-1999; | 99US-0155486. |
| PR | 08-JUL-1999; | 99US-0142803. | PR | 24-SEP-1999; | 99US-0155659. |
| PR | 09-JUL-1999; | 99US-0142920. | PR | 28-SEP-1999; | 99US-0156458. |
| PR | 12-JUL-1999; | 99US-0142977. | PR | 29-SEP-1999; | 99US-0156596. |
| PR | 13-JUL-1999; | 99US-0143542. | PR | 04-OCT-1999; | 99US-0157117. |
| PR | 14-JUL-1999; | 99US-0143624. | PR | 05-OCT-1999; | 99US-0157753. |
| PR | 15-JUL-1999; | 99US-0144005. | PR | 06-OCT-1999; | 99US-0157865. |
| PR | 16-JUL-1999; | 99US-0144085. | PR | 07-OCT-1999; | 99US-0158029. |
| PR | 16-JUL-1999; | 99US-0144086. | PR | 08-OCT-1999; | 99US-0158232. |
| PR | 19-JUL-1999; | 99US-0144325. | PR | 12-OCT-1999; | 99US-0158369. |
| PR | 19-JUL-1999; | 99US-0144331. | PR | 13-OCT-1999; | 99US-0159293. |
| PR | 19-JUL-1999; | 99US-0144332. | PR | 13-OCT-1999; | 99US-0159294. |
| PR | 19-JUL-1999; | 99US-0144333. | PR | 13-OCT-1999; | 99US-0159295. |
| PR | 19-JUL-1999; | 99US-0144334. | PR | 14-OCT-1999; | 99US-0159329. |
| PR | 19-JUL-1999; | 99US-0144335. | PR | 14-OCT-1999; | 99US-0159330. |
| PR | 20-JUL-1999; | 99US-0144352. | PR | 14-OCT-1999; | 99US-0159331. |
| PR | 20-JUL-1999; | 99US-0144632. | PR | 14-OCT-1999; | 99US-0159637. |
| PR | 20-JUL-1999; | 99US-0144684. | PR | 14-OCT-1999; | 99US-0159638. |
| PR | 21-JUL-1999; | 99US-0144814. | PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-JUL-1999; | 99US-0145086. | PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-JUL-1999; | 99US-0145088. | PR | 21-OCT-1999; | 99US-0160767. |
| PR | 22-JUL-1999; | 99US-0145085. | PR | 21-OCT-1999; | 99US-0160768. |
| PR | 22-JUL-1999; | 99US-0145087. | PR | 21-OCT-1999; | 99US-0160770. |
| PR | 22-JUL-1999; | 99US-0145089. | PR | 21-OCT-1999; | 99US-0160814. |
| PR | 22-JUL-1999; | 99US-0145192. | PR | 21-OCT-1999; | 99US-0160815. |
| PR | 23-JUL-1999; | 99US-0145145. | PR | 22-OCT-1999; | 99US-0160980. |
| PR | 23-JUL-1999; | 99US-0145218. | PR | 22-OCT-1999; | 99US-0160981. |
| PR | 23-JUL-1999; | 99US-0145224. | PR | 22-OCT-1999; | 99US-0160989. |
| PR | 26-JUL-1999; | 99US-0145276. | PR | 25-OCT-1999; | 99US-0161404. |
| PR | 27-JUL-1999; | 99US-0145913. | PR | 25-OCT-1999; | 99US-0161405. |
| PR | 27-JUL-1999; | 99US-0145918. | PR | 25-OCT-1999; | 99US-0161406. |
| PR | 27-JUL-1999; | 99US-0145919. | PR | 26-OCT-1999; | 99US-0161359. |
| PR | 28-JUL-1999; | 99US-0145951. | PR | 26-OCT-1999; | 99US-0161360. |
| PR | 02-AUG-1999; | 99US-0146386. | PR | 26-OCT-1999; | 99US-0161361. |
| PR | 02-AUG-1999; | 99US-0146388. | PR | 28-OCT-1999; | 99US-0161920. |
| PR | 02-AUG-1999; | 99US-0146389. | PR | 28-OCT-1999; | 99US-0161992. |
| PR | 03-AUG-1999; | 99US-0147038. | PR | 28-OCT-1999; | 99US-0161993. |
| PR | 04-AUG-1999; | 99US-0147204. | PR | 29-OCT-1999; | 99US-0162142. |
| PR | 04-AUG-1999; | 99US-0147302. |    |              |               |
| PR | 05-AUG-1999; | 99US-0147192. |    |              |               |
| PR | 05-AUG-1999; | 99US-0147260. |    |              |               |
| PR | 06-AUG-1999; | 99US-0147303. |    |              |               |
| PR | 06-AUG-1999; | 99US-0147416. |    |              |               |
| PR | 09-AUG-1999; | 99US-0147493. |    |              |               |
| PR | 09-AUG-1999; | 99US-0147935. |    |              |               |
| PR | 10-AUG-1999; | 99US-0148171. |    |              |               |
| PR | 11-AUG-1999; | 99US-0148319. |    |              |               |
| PR | 12-AUG-1999; | 99US-0148341. |    |              |               |
| PR | 13-AUG-1999; | 99US-0148565. |    |              |               |
| PR | 13-AUG-1999; | 99US-0148684. |    |              |               |
| PR | 16-AUG-1999; | 99US-0149368. |    |              |               |
| PR | 17-AUG-1999; | 99US-0149175. |    |              |               |
| PR | 18-AUG-1999; | 99US-0149426. |    |              |               |
| PR | 20-AUG-1999; | 99US-0149722. |    |              |               |
| PR | 20-AUG-1999; | 99US-0149723. |    |              |               |
| PR | 20-AUG-1999; | 99US-0149929. |    |              |               |
| PR | 23-AUG-1999; | 99US-0149902. |    |              |               |
| PR | 23-AUG-1999; | 99US-0149930. |    |              |               |
| PR | 25-AUG-1999; | 99US-0150566. |    |              |               |
| PR | 26-AUG-1999; | 99US-0150884. |    |              |               |
| PR | 27-AUG-1999; | 99US-0151065. |    |              |               |
| PR | 27-AUG-1999; | 99US-0151066. |    |              |               |
| PR | 27-AUG-1999; | 99US-0151080. |    |              |               |
| PR | 30-AUG-1999; | 99US-0151303. |    |              |               |
| PR | 31-AUG-1999; | 99US-0151438. |    |              |               |
| PR | 01-SEP-1999; | 99US-0151930. |    |              |               |
| PR | 07-SEP-1999; | 99US-0152363. |    |              |               |
| PR | 10-SEP-1999; | 99US-0153070. |    |              |               |
| PR | 13-SEP-1999; | 99US-0153758. |    |              |               |
| PR | 15-SEP-1999; | 99US-0154018. |    |              |               |

|                       |     |   |                  |                     |           |             |
|-----------------------|-----|---|------------------|---------------------|-----------|-------------|
| Query Match           |     |   | 39.2%;           | Score 1282;         | DB 21;    | Length 465; |
| Best Local Similarity |     |   | 62.3%;           | Pred. No. 1.5e-120; |           |             |
| Matches 243;          |     |   | Conservative 60; | Mismatches 81;      | Indels 6; | Gaps 5;     |
| Qy                    | 226 | ELFVPEKLDHVRAEAEPLSLISITKLDLQWVQLSEGWATPKGFMREKEYLQVMHPTD   | 285              |                     |           |             |
|                       |     | : |                  |                     |           |             |

RESULT 11  
AAG30252  
ID AAG30252 standard; Protein; 510 AA.  
XX AC AAG30252;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36134.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141642.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145219.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.



PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139432.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151085.  
PR 27-AUG-1999; 99US-0151086.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.

```

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      38.9%; Score 1272.5; DB 21; Length 473;
Best Local Similarity 57.0%; Pred. No. 1.4e-119;
Matches 243; Conservative 72; Mismatches 92; Indels 19; Gaps 6;

QY 207 VELLOQNVPIYTIK-----DIHFLVPENKLDHVRAEATLPSLSTIKD 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 39 LNLVYKRNLTMOSVSKMTVKSSLLDPDGGELVELIVPETEIGVKKRAESETMPKVLNQID 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 LQWQVLSGWTATPLKGFREKEYLQVMHFDL-LDDGV-INMSIPVLVPSAEDKTRLE 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 LEWVHVISEGWSPLKGFREDEYQLSHFNLSRLKNGTFVWNSLPVLAIDDDTKQET- 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 GCSKVF--LAHGSRVAILRDAAEFYEHKREERCSRWGTTCTKHPI-KMVMESGDWLG 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 GSENVALVCPGGDIIIGSLSRSEIYKHKEERARTWGTSPGLPVVEYITPSGNWLG 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 GDQVLEKIRWNGDLDQYRLTPLELKOKCKEMNADAVAFQLRNPVNHGHALLMDQTCRR 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 GDLVEFEPKYNDGLDHLRLSPKQREEDNRQADAVAFQLRNPVNHGHALLMDTRKR 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 LLERGKHPVLLHPLGGTCKDDVPLDRMKQHAHVLEEGVLDPKSTIVAIFFSPMLYA 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 278 LLEMGYKNVLLHPLGGTCKADDVPLDVRMEQHSKVLDDGVLDPKTTIVSIFPSPMHYA 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 489 GTEVQWHCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMSAPGLTSVEII 548
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 GTEVQWHAKARINAGANFYIVGRDPAGMHPGTEKRDLYDPDGHKRVLSMAFGLEKLNIL 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 549 PFRVAAYNAKKAMFYDPARINEFDIFSGTRMKLAREGENPPDGFMAPKAWKVLTDDY 608
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 PFRVAAYDTEIKMAFFDPSRAKEFLFISGTRMRTYARTGENPPDGFMCPSGWNVLVXY 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 609 RSEMDK 614
Db 458 ESLQEK 463

RESULT 13
AAG09594
ID AAG09594 standard; Protein; 476 AA.
XX AC AAG09594;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7589.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.

```

```

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

```



|    |              |               |
|----|--------------|---------------|
| PR | 20-JUL-1999; | 99US-0144332. |
| PR | 20-JUL-1999; | 99US-0144632. |
| PR | 20-JUL-1999; | 99US-0144884. |
| PR | 21-JUL-1999; | 99US-0144814. |
| PR | 21-JUL-1999; | 99US-0145086. |
| PR | 21-JUL-1999; | 99US-0145088. |
| PR | 22-JUL-1999; | 99US-0145085. |
| PR | 22-JUL-1999; | 99US-0145087. |
| PR | 22-JUL-1999; | 99US-0145089. |
| PR | 22-JUL-1999; | 99US-0145102. |
| PR | 23-JUL-1999; | 99US-0145145. |
| PR | 23-JUL-1999; | 99US-0145218. |
| PR | 23-JUL-1999; | 99US-0145224. |
| PR | 26-JUL-1999; | 99US-0145276. |
| PR | 27-JUL-1999; | 99US-0145913. |
| PR | 27-JUL-1999; | 99US-0145918. |
| PR | 27-JUL-1999; | 99US-0145919. |
| PR | 28-JUL-1999; | 99US-0145951. |
| PR | 02-AUG-1999; | 99US-0146386. |
| PR | 02-AUG-1999; | 99US-0146388. |
| PR | 02-AUG-1999; | 99US-0146389. |
| PR | 03-AUG-1999; | 99US-0147038. |
| PR | 04-AUG-1999; | 99US-0147204. |
| PR | 04-AUG-1999; | 99US-0147302. |
| PR | 05-AUG-1999; | 99US-0147192. |
| PR | 05-AUG-1999; | 99US-0147260. |
| PR | 06-AUG-1999; | 99US-0147303. |
| PR | 06-AUG-1999; | 99US-0147416. |
| PR | 09-AUG-1999; | 99US-0147493. |
| PR | 09-AUG-1999; | 99US-0147935. |
| PR | 10-AUG-1999; | 99US-0148171. |
| PR | 11-AUG-1999; | 99US-0148319. |
| PR | 12-AUG-1999; | 99US-0148341. |
| PR | 13-AUG-1999; | 99US-0148565. |
| PR | 13-AUG-1999; | 99US-0148684. |
| PR | 16-AUG-1999; | 99US-0149368. |
| PR | 17-AUG-1999; | 99US-0149175. |
| PR | 18-AUG-1999; | 99US-0149426. |
| PR | 20-AUG-1999; | 99US-0149722. |
| PR | 20-AUG-1999; | 99US-0149723. |
| PR | 23-AUG-1999; | 99US-0149929. |
| PR | 23-AUG-1999; | 99US-0149902. |
| PR | 23-AUG-1999; | 99US-0149930. |
| PR | 25-AUG-1999; | 99US-0150566. |
| PR | 26-AUG-1999; | 99US-0150884. |
| PR | 27-AUG-1999; | 99US-0151065. |
| PR | 27-AUG-1999; | 99US-0151066. |
| PR | 27-AUG-1999; | 99US-0151080. |
| PR | 30-AUG-1999; | 99US-0151303. |
| PR | 31-AUG-1999; | 99US-0151438. |
| PR | 01-SEP-1999; | 99US-0151930. |
| PR | 07-SEP-1999; | 99US-0152363. |
| PR | 10-SEP-1999; | 99US-0153070. |
| PR | 13-SEP-1999; | 99US-0153758. |
| PR | 15-SEP-1999; | 99US-0154018. |
| PR | 16-SEP-1999; | 99US-0154039. |
| PR | 20-SEP-1999; | 99US-0154779. |
| PR | 22-SEP-1999; | 99US-0155139. |
| PR | 23-SEP-1999; | 99US-0155486. |
| PR | 24-SEP-1999; | 99US-0155659. |
| PR | 28-SEP-1999; | 99US-0156458. |
| PR | 29-SEP-1999; | 99US-0156596. |
| PR | 04-OCT-1999; | 99US-0157117. |
| PR | 05-OCT-1999; | 99US-0157753. |
| PR | 06-OCT-1999; | 99US-0157865. |
| PR | 07-OCT-1999; | 99US-0158029. |
| PR | 08-OCT-1999; | 99US-0158232. |
| PR | 12-OCT-1999; | 99US-0158369. |
| PR | 13-OCT-1999; | 99US-0159293. |
| PR | 13-OCT-1999; | 99US-0159294. |
| PR | 13-OCT-1999; | 99US-0159295. |
| PR | 14-OCT-1999; | 99US-0159329. |
| PR | 14-OCT-1999; | 99US-0159330. |
| PR | 14-OCT-1999; | 99US-0159331. |
| PR | 14-OCT-1999; | 99US-0159637. |
| PR | 14-OCT-1999; | 99US-0159638. |
| PR | 18-OCT-1999; | 99US-0159584. |
| PR | 21-OCT-1999; | 99US-0160741. |
| PR | 21-OCT-1999; | 99US-0160767. |
| PR | 21-OCT-1999; | 99US-0160768. |
| PR | 21-OCT-1999; | 99US-0160770. |
| PR | 21-OCT-1999; | 99US-0160814. |
| PR | 21-OCT-1999; | 99US-0160815. |
| PR | 22-OCT-1999; | 99US-0160980. |
| PR | 22-OCT-1999; | 99US-0160981. |
| PR | 22-OCT-1999; | 99US-0160989. |
| PR | 25-OCT-1999; | 99US-0161404. |
| PR | 25-OCT-1999; | 99US-0161405. |
| PR | 25-OCT-1999; | 99US-0161406. |
| PR | 26-OCT-1999; | 99US-0161359. |
| PR | 26-OCT-1999; | 99US-0161360. |
| PR | 26-OCT-1999; | 99US-0161361. |
| PR | 28-OCT-1999; | 99US-0161920. |
| PR | 28-OCT-1999; | 99US-0161992. |
| PR | 28-OCT-1999; | 99US-0161993. |
| PR | 29-OCT-1999; | 99US-0162142. |

Query Match 38.98; Score 1272.5; DB 21; Length 476;

Best Local Similarity 57.08; Pred. No. 1.4e-119;

Matches 243; Conservative 72; Mismatches 92; Indels 19; Gaps 6;

|    |     |  |
|----|-----|--|
| Qy | 207 | VELLOEONIVPTIITK-----DIHELFVPEKNKLDHVRAEATLPSLSITKLD 253         |
| Db | 42  | LNLYKRNLTMQSVSKMTVTKSSLLDDPGGELVELIVPTEIGVKKAESETNPKVKLLNOLD 101 |
| Qy | 254 | LOWQVILSEGWATPLKGFMRKEKYLQVMHFDL-LDDGV-INMSIPVLPVSAEDKTRLE 311   |
| Db | 102 | LEWHVISEGWASPLKGFMRDEYLSLHFNLSLKNKGTFFVMSLPIVLAIDDDTKEQI- 160    |
| Qy | 312 | GCSKEV--LAHGGRRVAILRDAEYEHKRECRSVNGTCTCKPHI-KVMESGDWLVG 368      |
| Db | 161 | GSENVALVCPQGDIIIGLSRSVEIYKHNKEERARTWGTSPGLPYVEEYITPSCNWLIG 220   |
| Qy | 369 | GDLOVLEKIRWNGDLOYRLTPELEKOKCKEMNADAVAFOLRNPVHNGHALLMODTCRR 428   |
| Db | 221 | GDLEVFPIKINDGLDHYRUSPKOLREEDFNQADAVAFOLRNPVHNGHALLMNDTRKR 280    |
| Qy | 429 | LLERGYKHPVLLHPLGGWTKDDVPLDWRMKOHAAYLEEGVLPKSTIVAIFPSPMLYA 488    |
| Db | 281 | LLENGYKNPVLLHPLGGFTKADDVPLDVRMEQHSKVLDEGVLDPKTTIVSIFSPWHYA 340   |
| Qy | 489 | GPTVEQWCHRSRMITAGANFYIVGRDPAGMPHPETKKDLYEPHGGKVLSPAGLTSVEII 548  |
| Db | 341 | GPTVEQWHAARINAGANFYIVGRDPAGMGHPTEKRDLYDPDHGKRVLSMAGLEKNIL 400    |
| Qy | 549 | PFVAAYNKAKAMDYDPAHNEFDISGTRMKLAREGENPPDGFMAPKAWKVLTDYY 608       |
| Db | 401 | PFVAAYDTIEKNMAFFDPSRAKEFLISGTRMKTYARTGENPDPGFCPSGWNVLKYY 460     |
| Qy | 609 | RSEMDK 614   |
| Db | 461 | ESLQEK 466   |

RESULT 14

AAG09596

ID AAG09596 standard; Protein; 425 AA.

XX AC AAG09596;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7591.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

```

KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 18-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147933.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
```

|  |              |  |     |     |                                       |     |
|--|--------------|--|-----|-----|---------------------------------------|-----|
| PR   | 10-SEP-1999; | 99US-0153070.  | QY  | 578 | CTRMKRLAREGENPPDGFMAPKAWKVLTDIYKSEMDK | 614 |
| PR   | 13-SEP-1999; | 99US-0153758.  | Db  | 379 | GTKMTYARTGENPPDGFMCPSGWNVLVKYIESLOEK  | 415 |
| PR   | 15-SEP-1999; | 99US-0154018.  |     |     |                                       |     |
| PR   | 16-SEP-1999; | 99US-0154039.  |     |     |                                       |     |
| PR   | 20-SEP-1999; | 99US-0154779.  |     |     |                                       |     |
| PR   | 22-SEP-1999; | 99US-0155139.  |     |     |                                       |     |
| PR   | 23-SEP-1999; | 99US-0155486.  |     |     |                                       |     |
| PR   | 24-SEP-1999; | 99US-0155659.  |     |     |                                       |     |
| PR   | 28-SEP-1999; | 99US-0156458.  |     |     |                                       |     |
| PR   | 29-SEP-1999; | 99US-0156596.  |     |     |                                       |     |
| PR   | 04-OCT-1999; | 99US-0157117.  |     |     |                                       |     |
| PR   | 05-OCT-1999; | 99US-0157753.  |     |     |                                       |     |
| PR   | 06-OCT-1999; | 99US-0157863.  |     |     |                                       |     |
| PR   | 07-OCT-1999; | 99US-0158029.  |     |     |                                       |     |
| PR   | 08-OCT-1999; | 99US-0158232.  |     |     |                                       |     |
| PR   | 12-OCT-1999; | 99US-0158369.  |     |     |                                       |     |
| PR   | 13-OCT-1999; | 99US-0159293.  |     |     |                                       |     |
| PR   | 13-OCT-1999; | 99US-0159294.  |     |     |                                       |     |
| PR   | 13-OCT-1999; | 99US-0159295.  |     |     |                                       |     |
| PR   | 14-OCT-1999; | 99US-0159329.  |     |     |                                       |     |
| PR   | 14-OCT-1999; | 99US-0159330.  |     |     |                                       |     |
| PR   | 14-OCT-1999; | 99US-0159331.  |     |     |                                       |     |
| PR   | 14-OCT-1999; | 99US-0159637.  |     |     |                                       |     |
| PR   | 14-OCT-1999; | 99US-0159638.  |     |     |                                       |     |
| PR   | 18-OCT-1999; | 99US-0159584.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160741.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160767.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160768.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160770.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160814.  |     |     |                                       |     |
| PR   | 21-OCT-1999; | 99US-0160815.  |     |     |                                       |     |
| PR   | 22-OCT-1999; | 99US-0160980.  |     |     |                                       |     |
| PR   | 22-OCT-1999; | 99US-0160989.  |     |     |                                       |     |
| PR   | 22-OCT-1999; | 99US-0160989.  |     |     |                                       |     |
| PR   | 25-OCT-1999; | 99US-0161404.  |     |     |                                       |     |
| PR   | 25-OCT-1999; | 99US-0161405.  |     |     |                                       |     |
| PR   | 25-OCT-1999; | 99US-0161406.  |     |     |                                       |     |
| PR   | 26-OCT-1999; | 99US-0161359.  |     |     |                                       |     |
| PR   | 26-OCT-1999; | 99US-0161360.  |     |     |                                       |     |
| PR   | 26-OCT-1999; | 99US-0161361.  |     |     |                                       |     |
| PR   | 28-OCT-1999; | 99US-0161920.  |     |     |                                       |     |
| PR   | 28-OCT-1999; | 99US-0161922.  |     |     |                                       |     |
| PR   | 28-OCT-1999; | 99US-0161993.  |     |     |                                       |     |
| PR   | 29-OCT-1999; | 99US-0162142.  |     |     |                                       |     |
| Query Match 38.8%; Score 1270; DB 21; Length 425;              |              |  |     |     |                                       |     |
| Best Local Similarity 60.5%; Pred. No. 2.1e-119;               |              |  |     |     |                                       |     |
| Matches 240; Conservative 65; Mismatches 86; Indels 6; Gaps 5; |              |  |     |     |                                       |     |
| QY   | 223          | DIHELFPENKLDHVRAEATLPSLSITKLDLQWQVLSSEGWATPLKGFMRKEKYLQVMH | 282 |     |                                       |     |
| Db   | 20           | ELVELIVPETETGVKKAESMPKVKLNQIDLEWVHVISEGWSPLKGFMRDEYLSLH    | 79  |     |                                       |     |
| QY   | 283          | FDLL-LDDGV-INMSIPVLPSAEDKTRLEGCSKFV--LAHGGRRVAILRDAEFYHRK  | 338 |     |                                       |     |
| Db   | 80           | FNSLRKNGTFNMSPLVLAIDDDTKEQI-GSENVALVCPQGDIIIGLSRSEVIYKRNK  | 138 |     |                                       |     |
| QY   | 339          | EECRSRVGTCTTKPHI-KVMESGDWLVGDLQVLEKIRWNGDQYRLTLPLELKOC     | 397 |     |                                       |     |
| Db   | 139          | EERTARTWGTSPGLPYVEEYITPSGNWLGDLVFEPIKYNDGLDHYRLSPKQLREF    | 198 |     |                                       |     |
| QY   | 398          | KEMNADAVATQLRNPVINGHALLMODTCRLLERGYKHPVLLHPLGCGTKDDVPDLW   | 457 |     |                                       |     |
| Db   | 199          | DNROQADAVATQLRNPVINGHALLMODTCRLLERGYKHPVLLHPLGCGTKDDVPDLW  | 258 |     |                                       |     |
| QY   | 458          | RMKOHAAVLESGDLPKSTIVAIPEPSMLYAGPTEVQVHCRSMIAGANFYVGRDPAGM  | 517 |     |                                       |     |
| Db   | 259          | RMEQSKVLESGDLPKSTIVAIPEPSMLYAGPTEVQVHCRSMIAGANFYVGRDPAGM   | 318 |     |                                       |     |
| QY   | 518          | PHPEFKDLYEPTHGGKVLSPAGLTSVEIIPFRVAAYNKAAMDMFYDPAHNEFDNIS   | 577 |     |                                       |     |
| Db   | 319          | GHPTKRDLYDHDGKRVLSNAPGLEKNILPFRVAAYDTIEKKWAFDPSKAEFLIS     | 378 |     |                                       |     |

|           |  |                                       |     |
|-----------|--|---------------------------------------|-----|
| QY        | 578  | CTRMKRLAREGENPPDGFMAPKAWKVLTDIYKSEMDK | 614 |
| Db        | 379  | GTKMTYARTGENPPDGFMCPSGWNVLVKYIESLOEK  | 415 |
| RESULT 15 |  |                                       |     |
| ID        | AAG43667   |                                       |     |
| XX        | AAG43667 standard; Protein; 408 AA.                                      |                                       |     |
| AC        | AAG43667;  |                                       |     |
| XX        | 18-OCT-2000 (first entry)  |                                       |     |
| DT        | Arabidopsis thaliana protein fragment SEQ ID NO: 54606.                  |                                       |     |
| DE        | Protein identification; signal transduction pathway; metabolic pathway;  |                                       |     |
| XX        | hybridisation assay; genetic mapping; gene expression control; promoter; |                                       |     |
| KW        | termination sequence.  |                                       |     |
| XX        | Arabidopsis thaliana.  |                                       |     |
| OS        | EP1033405-A2.  |                                       |     |
| PN        | 06-SEP-2000.   |                                       |     |
| XX        | 25-FEB-2000; 2000EP-0301439.   |                                       |     |
| PF        | 99US-0121825.  |                                       |     |
| PR        | 99US-0123180.  |                                       |     |
| PR        | 99US-0123548.  |                                       |     |
| PR        | 99US-0125788.  |                                       |     |
| PR        | 99US-0126264.  |                                       |     |
| PR        | 99US-0126785.  |                                       |     |
| PR        | 99US-0127462.  |                                       |     |
| PR        | 99US-0128234.  |                                       |     |
| PR        | 99US-0128714.  |                                       |     |
| PR        | 99US-0129845.  |                                       |     |
| PR        | 99US-0130077.  |                                       |     |
| PR        | 99US-0130449.  |                                       |     |
| PR        | 99US-0130510.  |                                       |     |
| PR        | 99US-0130891.  |                                       |     |
| PR        | 99US-0131449.  |                                       |     |
| PR        | 99US-0132048.  |                                       |     |
| PR        | 99US-0132407.  |                                       |     |
| PR        | 99US-0132484.  |                                       |     |
| PR        | 99US-0132485.  |                                       |     |
| PR        | 99US-0132486.  |                                       |     |
| PR        | 99US-0132487.  |                                       |     |
| PR        | 99US-0132863.  |                                       |     |
| PR        | 99US-0134256.  |                                       |     |
| PR        | 99US-0134218.  |                                       |     |
| PR        | 99US-0134219.  |                                       |     |
| PR        | 99US-0134221.  |                                       |     |
| PR        | 99US-0134370.  |                                       |     |
| PR        | 99US-0134768.  |                                       |     |
| PR        | 99US-0134941.  |                                       |     |
| PR        | 99US-0135124.  |                                       |     |
| PR        | 99US-0135353.  |                                       |     |
| PR        | 99US-0135629.  |                                       |     |
| PR        | 99US-0136021.  |                                       |     |
| PR        | 99US-0136392.  |                                       |     |
| PR        | 99US-0136782.  |                                       |     |
| PR        | 99US-0137222.  |                                       |     |
| PR        | 99US-0137528.  |                                       |     |
| PR        | 99US-0137502.  |                                       |     |
| PR        | 99US-0137724.  |                                       |     |
| PR        | 99US-0138094.  |                                       |     |
| PR        | 99US-0138540.  |                                       |     |
| PR        | 99US-0138847.  |                                       |     |
| PR        | 99US-0139119.  |                                       |     |
| PR        | 99US-0139452.  |                                       |     |
| PR        | 99US-0139453.  |                                       |     |
| PR        | 99US-0139492.  |                                       |     |

```
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.5%; Score 1259.5; DB 21; Length 408;
Best Local Similarity 60.7%; Pred. No. 2.3e-118;
Matches 236; Conservative 69; Mismatches 79; Indels 5; Gaps 5;

QY 227 LFVPENKLDHIVRAEATLPS-LSITKLDLOWOVLSGWAATPLKGMREKEYLQVWVHDT 285
DB 3 LVVEESRRRVKWKHEATVTPARIKLNKRVLEWVLSGWAATPLKGMREKEYLQVWVHDT 62
QY 286 L-LDDG-VINMSTIPVLPVSAEDKTRLEGSCFVLAHG-GRRVAILDAEFYEHKREERC 342
DB 63 FRIEDGVSVMNSVPVILAIDDDDQKFRIGDSNQVTLVDVSGNPIALNDIIYKPKREERI 122
```

Search completed: May 30, 2003, 12:17:32  
Job time : 76 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:16:34 ; Search time 27 Seconds  
(without alignments)  
670.189 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSIGIKKQKTENQOKSTNVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No... is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description         |
|------------|--------|-------------|--------|-------|---------------------|
| 1          | 2602   | 79.5        | 624    | 2     | US-08-879-561-3     |
| 2          | 2599   | 79.5        | 624    | 2     | US-08-879-561-9     |
| 3          | 2367   | 72.4        | 610    | 2     | US-08-879-561-10    |
| 4          | 1274   | 38.9        | 463    | 4     | US-09-346-408-11    |
| 5          | 1268.5 | 38.8        | 476    | 4     | US-09-346-408-12    |
| 6          | 1263   | 38.6        | 465    | 4     | US-09-346-408-6     |
| 7          | 1237.5 | 37.8        | 465    | 4     | US-09-346-408-4     |
| 8          | 1227   | 37.5        | 461    | 4     | US-09-346-408-8     |
| 9          | 562    | 17.2        | 133    | 4     | US-09-346-408-2     |
| 10         | 481    | 14.7        | 204    | 4     | US-09-134-001C-4517 |
| 11         | 442.5  | 13.5        | 401    | 4     | US-09-134-001C-4403 |
| 12         | 417    | 12.7        | 110    | 4     | US-09-346-408-10    |
| 13         | 323    | 9.9         | 174    | 4     | US-09-149-476-635   |
| 14         | 198    | 6.1         | 36     | 4     | US-09-149-476-634   |
| 15         | 106    | 3.2         | 816    | 1     | US-08-038-760-3     |
| 16         | 106    | 3.2         | 816    | 2     | US-08-470-091-3     |
| 17         | 106    | 3.2         | 928    | 1     | US-08-204-329-1     |
| 18         | 106    | 3.2         | 928    | 2     | US-08-959-638-8     |
| 19         | 106    | 3.2         | 928    | 2     | US-08-482-627-5     |
| 20         | 106    | 3.2         | 928    | 3     | US-08-801-092-4     |
| 21         | 106    | 3.2         | 928    | 4     | US-08-328-673A-8    |
| 22         | 106    | 3.2         | 928    | 4     | US-09-315-113-4     |
| 23         | 106    | 3.2         | 928    | 5     | PCT-US94-10357-2    |
| 24         | 104.5  | 3.2         | 5588   | 4     | US-09-036-987A-6    |
| 25         | 104.5  | 3.2         | 5588   | 4     | US-09-370-700-6     |
| 26         | 104    | 3.2         | 928    | 5     | PCT-US94-10357-3    |
| 27         | 98.5   | 3.0         | 1365   | 6     | 5194600-4           |

Patent No. 5194600

ALIGNMENTS

RESULT 1

US-08-879-561-3

; Sequence 3, Application US/08879561

; Patent No. 5817482

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Hawkins, Phillip R.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/879,561

; FILING DATE: Herewith

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0325 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 624 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: LUNGNOR02

; CLONE: 373887

; US-08-879-561-3

Sequence 5377, Ap  
Sequence 29, Appl  
Sequence 29, Appl  
Sequence 30, Appl  
Sequence 30, Appl  
Sequence 46, Appl  
Sequence 18, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 28, Appl  
Sequence 28, Appl  
Sequence 5, Appl  
Sequence 4, Appl  
Sequence 119, App  
Sequence 42, Appl

Query Match 79.5%; Score 2602; DB 2; Length 624;  
Best Local Similarity 77.1%; Pred. No. 6.1e-269;  
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKQKTEN-----QKSTNVVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAKTTI 58  
DB 9 KVKLSNNAQNMGMQRATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAKTTV 68

QY 59 SFALSYLYSHATPCYSLDGDNVVRHGNLNLGSPGDRENIIRIAEVAKLADAGLVCI 118  
DB 69 SNALEYLVCHGIPCYLTLDGDNIRQGLNKLNGSPEDRENVIRIAEVAKLADAGLVCI 128

QY 119 TSFISPAKDRENAKIHESAGLPFFEIFVDAPLNICESRDVKGKLYKARAGIKGFTGI 178  
DB 129 TSFISPYTDORNNARQIHESAGLPFFEIFVDAPLNICESRDVKGKLYKARAGIKGFTGI 188

QY 179 DSDYKPEPTEPVKLTNLSTVSCVHQVVELLQEQNIPTIITKDIHELFPVKNKLDHVR 238  
DB 189 DSEYKPEAPELVKLTDCDNDVDCVQVVELLQEQNIPTIITKDIHELFPVKNKLDHVR 248

QY 239 AEATLPSLSITKLDLQWVQLSEGWATPLKGFMRKEYLQVNHFTDLDGGINNSIPI 298  
DB 309 VLTATHEDKERLDGCTAFALVYEGRRVAILRNPPEFHEHKEERCARQWGTCKNHPYIKM 368

QY 359 VMEGDMVLGGDLQVLEKIRWNDGLDQYRLTPELKQCKEMNADAVFAFOLRNPVHNGH 418  
DB 369 VMEGDMVLGGDLQVLDRIYVNDGLDQYRLTPELKQCKEMNADAVFAFOLRNPVHNGH 428

QY 419 ALLMQDTCRLLERGGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIV 478  
DB 429 ALLMQDTHQLLERGGYRVRPVLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGTPEVOWHCRSMIAGANFYIVGRDPAGMPHETKDKDLEPHGKVLMS 538  
DB 489 AIFPSPMYAGTPEVOWHCRSMIAGANFYIVGRDPAGMPHETKDKDLEPHGKVLMS 548

QY 539 APGLTSVEIIPFVAAYNAKAKAMDYDPARHNEFDIFSGTRMKRLAREGENPPDGFMAP 598  
DB 549 APGLTLEIVPFRVAAYNAKAKAMDYDPARHNEFDIFSGTRMKRLAREGENPPDGFMAP 608

QY 599 KAWKVLTDYRS 610  
DB 609 KAWTVLVEYKS 620

RESULT 2

US-08-879-561-9  
Sequence 9, Application US/08879561  
Patent No. 5817482  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Hawkins, Phillip R.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/879,561  
FILING DATE: Herewith  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0325 US  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1109876  
US-08-879-561-9

Query Match 79.5%; Score 2599; DB 2; Length 624;  
Best Local Similarity 77.0%; Pred. No. 1.3e-268;  
Matches 471; Conservative 66; Mismatches 69; Indels 6; Gaps 1;

QY 5 KKQKTEN-----QKSTNVVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAKTTI 58  
DB 9 KVKLSNNAQNMGMQRATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAKTTV 68

QY 59 SFALSYLYSHATPCYSLDGDNVVRHGNLNLGSPGDRENIIRIAEVAKLADAGLVCI 118  
DB 69 SNALEYLVCHGIPCYLTLDGDNIRQGLNKLNGSPEDRENVIRIAEVAKLADAGLVCI 128

QY 119 TSFISPAKDRENAKIHESAGLPFFEIFVDAPLNICESRDVKGKLYKARAGIKGFTGI 178  
DB 129 TSFISPYTDORNNARQIHESAGLPFFEIFVDAPLNICESRDVKGKLYKARAGIKGFTGI 188

QY 179 DSDYKPEPTEPVKLTNLSTVSCVHQVVELLQEQNIPTIITKDIHELFPVKNKLDHVR 238  
DB 189 DSEYKPEAPELVKLTDCDNDVDCVQVVELLQEQNIPTIITKDIHELFPVKNKLDHVR 248

QY 239 AEATLPSLSITKLDLQWVQLSEGWATPLKGFMRKEYLQVNHFTDLDGGINNSIPI 298  
DB 249 TDAEALPALKINKYDMQVQLAEGWATPLNGFMREREYLQCLHFDCLDGGVINSVPI 308

QY 299 VLPVSAEDKTRLEGGCKFVLAHGGRRVAILRDAEFYEHKEERCRCRVWGTCTCKHPIKIM 358  
DB 309 VLTATHEDKERLDGCTAFALVYEGRRVAILRNPPEFHEHKEERCARQWGTCKNHPYIKM 368

QY 359 VMEGDMVLGGDLQVLEKIRWNDGLDQYRLTPELKQCKEMNADAVFAFOLRNPVHNGH 418  
DB 369 VMEGDMVLGGDLQVLDRIYVNDGLDQYRLTPELKQCKEMNADAVFAFOLRNPVHNGH 428

QY 419 ALLMQDTCRLLERGGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIV 478  
DB 429 ALLMQDTHQLLERGGYRVRPVLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGTPEVOWHCRSMIAGANFYIVGRDPAGMPHETKDKDLEPHGKVLMS 538  
DB 489 AIFPSPMYAGTPEVOWHCRSMIAGANFYIVGRDPAGMPHETKDKDLEPHGKVLMS 548

QY 539 APGLTSVEIIPFVAAYNAKAKAMDYDPARHNEFDIFSGTRMKRLAREGENPPDGFMAP 598  
DB 549 APGLTLEIVPFRVAAYNAKAKAMDYDPARHNEFDIFSGTRMKRLAREGENPPDGFMAP 608

QY 599 KAWKVLTDYRS 610  
DB 609 KAWTVLVEYKS 620



RESULT 3  
US-08-879-561-10  
; Sequence 10, Application US/08879561  
; Patent No. 5817482  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Hawkins, Phillip R.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/879,561  
; FILING DATE: Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0325 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 705385  
US-08-879-561-10

Query Match 72.4%; Score 2367; DB 2; Length 610;  
Best Local Similarity 71.4%; Pred. No. 7,7e-244;  
Matches 432; Conservative 69; Mismatches 104; Indels 0; Gaps 0;  
QY 11 NOQSTNVYQAHVSRNKGQVGTGCGFRGCTVWLTGLSGAGKTTISFALEEYLVSHA 70  
DB 6 NGLATNVTFTQHVSRKRGQVLTGQGGFRGCTVWFTGLSGAGKTTISFALEEYLVSG 65  
QY 71 IPCYSLDGNVHGLNRLNGLSPGCDREENIRIAEVAKLFADAGLVCTISFPKADRE 130  
DB 66 IPTYSLDGNVHGLNRLNGLFTOEDREENIRISAEVAKLFADAGLVCTISFPKADRE 125  
QY 131 NARKTHESAGLPFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGDSYKPEPTER 190  
DB 126 LARSLHEQAGLPFFECFVDTPLDVCEQDRVKGLYKARAGIKGFTGIDQOYESDAPET 185  
QY 191 VLKTNLSVDCVHQVVELLOEQNIVPTTIKDIHELVPEPNKLDHVRAEATLPSLSIT 250  
DB 186 QLYAGNKSIDECQVBSLLQNGVVPESAVNIVKELFVPSGLEHAKEAIVDLPTMEIT 245  
QY 251 KLDLQWQVLSSEGWATPLKGFMRKEEYLVQVHFDLTDGGINNSIPVLPVSAEDKTRL 310

DB 246 KLDLQWQVLSSEGWATPLTGFMREREYLSQSHFGCLLDGGVTNQSIPIVLPVHTADKRL 305  
QY 311 EGCSEFVLAHGGRVAILRDAEFYHRKEERCSRVMGTCTTKHPHKKMYMESDGLVGGD 370  
DB 306 EGSSAFALSYEGKRIAILRTPEFYHRKEERCSRQFGTSNACQPVYVKMIMESDGLVGGD 365  
QY 371 LOVLEKIRWNDGLDQYRLTPELKKQCKCKEMNADAVFAFQLRNPVHNGHALLMQDTCRLL 430  
DB 366 LEVLERITWNDGLDEYRLTPNELRAKRALNADAVFAFQLRNPVHNGHALLMQDTCRLL 425  
QY 431 ERGYKHPVLLHLLPLGCGTKDDVPLDWRMKQHAANLEBGLVDPKSTIVAIFSPMLYAGP 490  
DB 426 ERGYKHPVLLHLLPLGCGTKDDVPLDWRMKQHAANLEBGLVDPKSTIVAIFSPMLYAGP 485  
QY 491 TEVQWCHRSRMIAAGNFYIVGRDPAGMHPETKDLIYETHGKVLSPAGLTSVPIPF 550  
DB 486 TEVQWCHRSRMIAAGNFYIVGRDPAGMHPETKDLIYETHGKVLSPAGLTSVPIPF 545  
QY 551 RVAAYNKAAMDFYDPAHNEFDISGTRMKLAREGNPPDGFMAPKAWKVLTDYYS 610  
DB 546 RVAAYNKAAMDFYDPAHNEFDISGTRMKLAREGNPPDGFMAPKAWKVLTDYYS 605  
QY 611 EMDKN 615  
DB 606 KAOQS 610

RESULT 4  
US-09-346-408-11  
; Sequence 11, Application US/09346408B  
; Patent No. 6338966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01  
; EARLIER APPLICATION NUMBER: 60/092,833  
; EARLIER FILING DATE: July 14, 1998  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 11  
; LENGTH: 463  
; TYPE: PRT  
; ORGANISM: Solanum tuberosum  
US-09-346-408-11

Query Match 38.9%; Score 1274; DB 4; Length 463;  
Best Local Similarity 61.4%; Pred. No. 3.2e-127;  
Matches 239; Conservative 60; Mismatches 86; Indels 4; Gaps 4;  
QY 226 ELFVPEKLDHVRAEATLPSLSITKLDQWQVLSSEGWATPLKGFMRKEEYLVQVHFD 285  
DB 60 ELIVEEPQDLKRRQALSLPQIKLSKIDIQWVHLSSEGWASPLKGFMRSEFQLTLHFS 119  
QY 286 L-LDQG-VINMSIPVLPVSAEDKTRLECGSKFVLAHG-GRVAILRDAEFYHRKEERC 342  
DB 120 LRIGGSSVVMSPVILVLAIDDSNNKNNIGDSSSVVALVDKDNPIALINDIEIKHKEERT 179  
QY 343 SRVWCTTCTKHPHKKM-VMESDGLVGGDLQVLEKIRWNDGLDQYRLTPELKKQCKEMN 401  
DB 180 ARTWCTTAGPLPYAQAIATHAGNWLIGDLEVIETPKYHGDGLDRFLSPALRDEFTRN 239  
QY 402 ADAFAFQLRNPVHNGHALLMQDTCRLLERGYKHPVLLHLLPLGCGTKDDVPLDWRMKQ 461  
DB 240 ADAFAFQLRNPVHNGHALLMQDTCRLLERGYKHPVLLHLLPLGCGTKDDVPLDWRMKQ 299  
QY 462 HAAVLEEGVLPKSTIVAIFSPMLYAGPTEVQWCHRSRMIAAGNFYIVGRDPAGMHPPE 521

```
Db 300 HEMVLEDCVLDPETTVVSIIPSPMHYACPTVEQVHAKARINAGANFYIVGRDPAGMGHPL 359
Qy 522 TKKDLYEPTGGKVL$MAPGLTSVEIIPFRVAAYNKAAMDFYDPAHNEFFDISGTRM 581
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 360 EKRDLYDADHGKVL$MAPGLERLNLPRVAAYDKTQNGMAFFDPSRQDFLISGTRM 419
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 582 RKLAREGENPPDGFMAPKANKVLTDDYYS 610
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 420 RALAKNKENPPDGFMCPPGGKVLVDYDYS 448
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

RESULT 5
US-09-346-408-12
; Sequence 12, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-346-408-12

Query Match 38.8%; Score 1268.5; DB 4; Length 476;
Best Local Similarity 57.3%; Pred. No. 1.3e-126;
Matches 242; Conservative 71; Mismatches 90; Indels 19; Gaps 6;

Qy 207 VELLQGNIVPYIIRK-----DIHELFPENKLDHVRAEATLPSLSTKLD 253
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 42 LNLVYRNLTMQSVKMTVKSSLDPDGGLVELIIVPETEIGVKKAESETMPKVKLNQID 101
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 254 LQWQVLSSEGATPLKGFMEKEVYLVHFDTL-LDDGV-INMSIPVLVPSAEDKTRLE 311
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 102 LEWVHWISEGASPLKGFMEDEYLOSLSHNSRLKNGTFVNSLPIVLAIIDDTKEQI- 160
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 312 GCSKFV--LAHGRRVAILRDAEFYEHKRECSRVMGTCTCKPHI-KMVMSGDWLVG 368
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 161 GSSENVALVCPQDIIIGSLRSVEIYKHKEERARTWGTSPGLPYVEEYITPSGNWLI 220
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 369 GDLOVLEKIRWNGLDQYRLTPLELKOKCKEMNADAVAFQLRNPVHNGHALLMQDTCRR 428
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 221 GDLEVFETIKYNDGLDHYRLSPKQLEEDFNQADAVAFQLRNPVHNGHALLMNDTRKR 280
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 429 LLERGKHPVLLHPLGGWTKDDVPLDWRMKAHVALEBGLDPKSTIVAIPSPMLYA 488
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 281 LLEMGYNPVLHPLGGFKADDVPLDVRMEGSKVLEGDVLDPKTIVSIIPSPMHYA 340
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 489 GPTEVQWHCSRMIAGANFYIVGRDPAGMPHPTKKDLYEPHGGKVL$MAPGLTSVEII 548
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 341 GPTEVQWHAKARINAGANFYIVGRDPAGMGHPTEKRDLYDPDHGKVL$MAPGLEKLNIL 400
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 549 PFRVAAYNKAAMDFYDPAHNEFFDISGTRMRKLAREGENPPDGFMAPKANKVLTDDY 608
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 401 PFRVAAYDTIEKKMAFFDPSRAKEFLISGTRMRTYARTGENPPDGFMCPSGNWLVKY 460
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 609 RS 610
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 461 ES 462

RESULT 6
US-09-346-408-6
; Sequence 4, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 465
```

```
; Sequence 6, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Glycine max
US-09-346-408-6

Query Match 38.6%; Score 1263; DB 4; Length 465;
Best Local Similarity 61.7%; Pred. No. 4.8e-126;
Matches 240; Conservative 59; Mismatches 86; Indels 4; Gaps 4;

Qy 226 ELFVPEKNKLDHVRAEATLPSLSTIKLDLQWQVLSSEGATPLKGFMEKEVYLVHFD 285
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 62 ELVVTDFERDLKKGALSLPRIKLSRIDLEWHVLSSEGATPLKGFMEKEVYLVHFD 121
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 286 L-LDDG-VINMSIPVLVPSAEDKTRLEGC$KFLAHG-GRRVAILRDAEFYEHKREERC 342
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 122 LRLDGSVVMNSVPIVLAIIDDAQKHRIKGNKVALFDSKGDPAVAILNIIYKPKKEERI 181
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 343 SRVWGTCTCKPHIKM-VMSGDWLVGGDLQVLEKIRWNGLDQYRLTPLELKOKCKEMN 401
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 182 ARTWGTIAPLTVEQITTNAGNWLIGDLEVEPIQYNDGLDHFRL$PAQLRAEFTTRN 241
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 402 ADAVAFQLRNPVHNGHALLMQDTCRLLERGKHPVLLHPLGGWTKDDVPLDWRMKA 461
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 242 ADAVAFQLRNPVHNGHALLMTDTRKRLLEMGYNPVLHPLGGYTKADDVPLDWRMKA 301
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 462 HAAVLEBGLDPKSTIVAIPSPMLYACPTVEQVHCSRMIAGANFYIVGRDPAGMPHPE 521
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 302 HEVLEDCVLDPETTVVSIIPSPMHYACPTVEQVHAKARINAGANFYIVGRDPAGMHPV 361
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 522 TKKDLYEPTGGKVL$MAPGLTSVEIIPFRVAAYNKAAMDFYDPAHNEFFDISGTRM 581
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 362 EKRDLYDADHGKVL$MAPGLERLNLPRVAAYDKTQNGMAFFDPSRQDFLISGTRM 421
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Qy 582 RKLAREGENPPDGFMAPKANKVLTDDYYS 610
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db 422 RTLARNKESPPDGFMCPPGGKVLVDYDYS 450
    ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

RESULT 7
US-09-346-408-4
; Sequence 4, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 465
```

226 ELFPENKLDHVRAEATLPSLSSTKLDLQWQVLSGEGWATPLKGFMRREKEYLQVMHFDT 285  
; Patent NO. 03603370  
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

SEQUENCE 4403, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT FILING DATE: 1998-08-13  
PRIOR FILING DATE: 1997-11-08  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4517  
LENGTH: 204  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4517

Query Match 14.7%; Score 481; DB 4; Length 204;  
Best Local Similarity 44.5%; Pred. No. 4.5e-43;  
Matches 93; Conservative 37; Mismatches 73; Indels 6; Gaps 2;

QY 8 KTNQOKSTNVVQAAHHSVRNKRQVVGTRGGRGCTVWLTGLSGAGKTTISFALBEYLV 67  
DB 2 RSVKMSSENHITWHDSEVTKQRQ-----KNGHKSAVINFTGLSGSGKSTVSVALEKELF 57  
QY 68 SHAIPCYSLDGNVHGLNRLNGLFSGPDREENIRIRAEVAKLADAGLVCIITSFISPPAK 127  
DB 58 NEGQTYRLDGNVHGLNRLNGLFSGPDREENIRIRAEVAKLADAGLVCIITSFISPPAK 117  
QY 128 DRGNARKIHESAGLPFFEIFVDAPLNICESRDVKGLYKRRAGEIKGTGIDSDYKPKET 167  
DB 118 DREGVRLLEDN--EFIEVYTKSVCECEKRDPKGLYKRRAGEIKGTGIDSDYKPKET 175  
QY 188 PERVLKNTSTVSDCVHQVVELLOEQNV 216  
DB 176 PEITIDTEHDTIEQSVVQIIRLKEHEYI 204

RESULT 11  
US-09-134-001C-4403  
Sequence 4403, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT FILING DATE: 1998-08-13  
PRIOR FILING DATE: 1997-11-08  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4403  
LENGTH: 401  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4403

Query Match 13.5%; Score 442.5; DB 4; Length 401;  
Best Local Similarity 29.9%; Pred. No. 1.9e-38;  
Matches 122; Conservative 82; Mismatches 175; Indels 29; Gaps 11;

QY 206 VVELLOEQNVYTIKDIHETL---VPEKLDHVAEAEITLPSITKLDLQWQVLS 262  
DB 7 LIIMSNNETITNTKPHGGLINRVVEGNERLIEEALKFKPTLNPWGISDLELIGI 66  
QY 263 GNATPLKGFMRKEYLQVHFDITLDDGVNINSIPVLVPSAEDKTRLE-GCSKFVLAHG 321  
DB 67 GGFSPUTGFMNKEDYTKVIE-ETHLSNGLV-WSIPTLVTESEADKLEIGDIALYGBD 124  
QY 322 GRRVAILRDAEYFHRKERCRCRVWGTTCTKPHIKMVMESGDLVGGDLQVLEKIRWD 381

Db 125 GOLYGTILKEEYTKYDKEEARLVYGTTEBAPGVKVEKGNILGGPIKLLNRPK-HD 183  
QY 382 GLDOYRLTLPLEKQCKEMNADAVAFQRLNPNVHNGHALLMODTCRLLRGGYKHPVLLL 441  
Db 184 AFSNYHLDPSSETQOLFDHGLGKNTVVGFTQTRNPVHRAHEYI-OXSALDIVDG-----LLL 236  
QY 442 HPLGGWTKDDVPLDMRMKQAAVLEEGVLDPKASTIVAIFPSPMLYAGPTEVQWHCGRM 501  
Db 237 NPLVGETKSDIPADVRMESYEVLKKNYPEDRARLV-IYPAARMYAGPREAILHATVRK 295  
QY 502 IAGANFYIVGRDPAGMPHPETKDKLDYEPHGGKVLSSMAPGLTSVEIIPFVAAV-----N 556  
Db 296 NYGCTHFIIVGRDHAGV-----GDYGTVEAQLITOFDELGTQILKFEHAFYCEACGN 349  
QY 557 KAKKAMDYDPARHNEFDTSGTRMRKLAREGENPPDGENAPKAWKVL 604  
Db 350 MATAKTCPHDASOHLH---LSGKVRKRLNGESLPTKFSRPEVAEVL 394

## RESULT 12

US-09-346-408-10  
Sequence 10, Application US/09346408B  
Patent No. 6338966  
GENERAL INFORMATION:  
APPLICANT: Allen, Steve  
APPLICANT: Anderson, Shawn  
APPLICANT: Falco, Carl  
APPLICANT: Rafalski, Antoni  
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
FILE REFERENCE: BB-1167-A  
CURRENT APPLICATION NUMBER: US/09/346,408B  
CURRENT FILING DATE: 1999-07-01  
EARLIER APPLICATION NUMBER: 60/092,833  
EARLIER FILING DATE: July 14, 1998  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 10  
LENGTH: 110  
TYPE: PRT  
ORGANISM: Triticum aestivum  
US-09-346-408-10

Query Match 12.7%; Score 417; DB 4; Length 110;  
Best Local Similarity 71.3%; Pred. No. 1.1e-36;  
Matches 77; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 359 VMESGDLVGGDLQVLEKIRWNDGLQYRLTLPLEKQCKEMNADAVAFQRLNPNVHNGH 418  
Db 3 ITNAGDLVGGDLQVLEKIRWNDGLQYRLTLPLEKQCKEMNADAVAFQRLNPNVHNGH 62  
QY 419 ALLMODTCRLLRGGYKHPVLLLHPLGGWTKDDVPLDMRMKQAAV 466  
Db 63 ALLMTDTRRLLEMGYKHPVLLLHPLGGWTKDDVPLDMRMKQAAV 110

## RESULT 13

US-09-149-476-635  
Sequence 635, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621

[illegible]

|  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |  |                                 |
|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|
| EARLIER APPLICATION NUMBER: 60/043,672 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/043,315 | EARLIER FILING DATE: 1997-04-11 | EARLIER APPLICATION NUMBER: 60/048,974 | EARLIER FILING DATE: 1997-06-06 | EARLIER APPLICATION NUMBER: 60/056,886 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,877 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,889 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,893 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,630 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,878 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,662 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,872 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,882 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,637 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,903 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,888 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,879 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,880 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,894 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,911 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,636 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,874 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,910 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,864 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,631 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,845 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/056,892 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/057,761 | EARLIER FILING DATE: 1997-08-22 | EARLIER APPLICATION NUMBER: 60/047,595 | EARLIER FILING DATE: 1997-03-23 | EARLIER APPLICATION NUMBER: 60/047,599 | EARLIER FILING DATE: 1997-03-23 | EARLIER APPLICATION NUMBER: 60/047,588 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,585 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,586 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,590 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,594 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,589 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,593 | EARLIER FILING DATE: 1997-05-23 | EARLIER APPLICATION NUMBER: 60/047,614 | EARLIER FILING DATE: 1997-05-23 |
|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|--|---------------------------------|

```

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match          9.9%; Score 323; DB 4; Length 174;
Best Local Similarity 83.3%; Pred. No. 2,6e-26;
Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy      1 MSGIKKOKTENQOKSTNVVQAHVSRNKRQGVVGRGTRGCTVWLTGLSGAGKTTISF 60
        |||||||
Db      55 MSGIKKOKTENQOKSTNVVQAHVSRNKRQGVVGRGTRGCTVWLTGLSGAGKNNDF 114
        |||||||

Qy      61 ALREYLVSHAIPCVSLDG 78
        |||||||
Db      115 CPGGVLYSHAIPVNSWNG 132

RESULT 14
US-09-149-476-634
; Sequence 634, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen-et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002p1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
```

; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/048,974  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/056,886  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,877  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,889  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,893  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,630  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,878  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,662  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,872  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,882  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,637  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,903  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,888  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,879  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,880  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,894  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,911  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,636  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,874  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,910  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,864  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,631  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,845  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,892  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,761  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/047,595  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,599  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,588  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,585  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,586  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,590  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,594  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,589  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,593  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,614  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,578  
 ; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,576  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/047,501  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,670  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/056,632  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,664  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,876  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,881  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,909  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,875  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,862  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,887  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,908  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/048,964  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/057,650  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/056,884  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,669  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/049,610  
 ; EARLIER FILING DATE: 1997-06-13  
 ; EARLIER APPLICATION NUMBER: 60/061,060  
 ; EARLIER FILING DATE: 1997-10-02

Query Match 6.1%; Score 198; DB 4; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 4e-14;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQVVGTRGFRGCVVLTGLSGACK 55

DB 1 YQAHVSRNKRQVVGTRGFRGCVVLTGLSGACK 36

RESULT 15

US-08-038-760-3

; Sequence 3, Application US/08038760

; Patent No. 5496731

; GENERAL INFORMATION:

; APPLICANT: Xu, Hong-Ji

; APPLICANT: Hu, Shi-Xue

; APPLICANT: Benedict, William F.

; TITLE OF INVENTION: Broad-Spectrum Tumor Suppressor Genes, Gene Products and Methods for Tumor Suppressor Gene Therapy.

; TITLE OF INVENTION: Methods for Tumor Suppressor Gene Therapy.

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/038,760

; FILING DATE: 19930325

; CLASSIFICATION: 424

US-08-038-760-3

729 TSEKFQKI-----NQM-VCNSDRVLKRS AEGSNPP 757



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:19:45 ; Search time 48 Seconds  
(without alignments)  
1296.929 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 3271  
Sequence: 1 MSGIKKQKTEKQKSTNVVY.....MAPKAWKLVTDYRSEMDKN 615

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_RA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Match | Length | ID | Description          |
|------------|-------|-------|--------|----|----------------------|
| 1          | 3216  | 98.3  | 619    | 9  | US-09-974-298-184    |
| 2          | 2602  | 79.5  | 624    | 9  | US-09-984-245-196    |
| 3          | 2602  | 79.5  | 624    | 9  | US-09-966-262-196    |
| 4          | 2602  | 79.5  | 624    | 9  | US-09-983-966-196    |
| 5          | 2602  | 79.5  | 624    | 9  | US-10-143-090-196    |
| 6          | 2588  | 79.1  | 624    | 10 | US-09-771-161A-244   |
| 7          | 571.5 | 17.5  | 162    | 9  | US-10-156-239-40     |
| 8          | 571.5 | 17.5  | 162    | 9  | US-10-199-485-40     |
| 9          | 571.5 | 17.5  | 162    | 10 | US-09-795-693-40     |
| 10         | 476.5 | 14.6  | 201    | 10 | US-09-741-669-457    |
| 11         | 340   | 10.4  | 65     | 10 | US-09-864-761-39981. |
| 12         | 323   | 9.9   | 174    | 9  | US-09-809-391-635    |
| 13         | 221.5 | 6.8   | 163    | 10 | US-09-867-550-96     |
| 14         | 217   | 6.6   | 41     | 9  | US-09-984-245-340    |
| 15         | 217   | 6.6   | 41     | 9  | US-09-966-262-340    |
| 16         | 217   | 6.6   | 41     | 9  | US-09-983-966-340    |
| 17         | 217   | 6.6   | 41     | 9  | US-10-143-090-340    |
| 18         | 198   | 6.1   | 36     | 9  | US-09-809-391-634    |
| 19         | 187   | 5.7   | 50     | 10 | US-09-771-161A-153   |

|    |     |     |    |   |                   |                   |
|----|-----|-----|----|---|-------------------|-------------------|
| 20 | 181 | 5.5 | 45 | 9 | US-09-984-245-332 | Sequence 332, App |
| 21 | 181 | 5.5 | 45 | 9 | US-09-966-262-332 | Sequence 332, App |
| 22 | 181 | 5.5 | 45 | 9 | US-09-983-966-332 | Sequence 332, App |
| 23 | 181 | 5.5 | 45 | 9 | US-10-143-090-332 | Sequence 332, App |
| 24 | 177 | 5.4 | 37 | 9 | US-09-984-245-339 | Sequence 339, App |
| 25 | 177 | 5.4 | 37 | 9 | US-09-966-262-339 | Sequence 339, App |
| 26 | 177 | 5.4 | 37 | 9 | US-09-983-966-339 | Sequence 339, App |
| 27 | 177 | 5.4 | 37 | 9 | US-10-143-090-339 | Sequence 339, App |
| 28 | 175 | 5.4 | 41 | 9 | US-09-984-245-336 | Sequence 336, App |
| 29 | 175 | 5.4 | 41 | 9 | US-09-966-262-336 | Sequence 336, App |
| 30 | 175 | 5.4 | 41 | 9 | US-09-983-966-336 | Sequence 336, App |
| 31 | 175 | 5.4 | 41 | 9 | US-10-143-090-336 | Sequence 336, App |
| 32 | 172 | 5.3 | 39 | 9 | US-09-984-245-338 | Sequence 338, App |
| 33 | 172 | 5.3 | 39 | 9 | US-09-966-262-338 | Sequence 338, App |
| 34 | 172 | 5.3 | 39 | 9 | US-09-983-966-338 | Sequence 338, App |
| 35 | 172 | 5.3 | 39 | 9 | US-10-143-090-338 | Sequence 338, App |
| 36 | 166 | 5.1 | 38 | 9 | US-09-984-245-333 | Sequence 333, App |
| 37 | 166 | 5.1 | 38 | 9 | US-09-966-262-333 | Sequence 333, App |
| 38 | 166 | 5.1 | 38 | 9 | US-09-983-966-333 | Sequence 333, App |
| 39 | 166 | 5.1 | 38 | 9 | US-10-143-090-333 | Sequence 333, App |
| 40 | 166 | 5.1 | 39 | 9 | US-09-984-245-334 | Sequence 334, App |
| 41 | 166 | 5.1 | 39 | 9 | US-09-966-262-334 | Sequence 334, App |
| 42 | 166 | 5.1 | 39 | 9 | US-09-983-966-334 | Sequence 334, App |
| 43 | 166 | 5.1 | 39 | 9 | US-10-143-090-334 | Sequence 334, App |
| 44 | 149 | 4.6 | 40 | 9 | US-09-984-245-335 | Sequence 335, App |
| 45 | 149 | 4.6 | 40 | 9 | US-09-966-262-335 | Sequence 335, App |

## ALIGNMENTS

RESULT 1  
US-09-974-298-184  
; Sequence 184, Application US/09974298  
; Patent No. US20020156263A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hwei-Mei  
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER  
; FILE REFERENCE: PA-0037 P  
; CURRENT APPLICATION NUMBER: US/09/974,298  
; PRIORITY FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 60/238,331  
; PRIOR FILING DATE: 2000-05-10  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 184  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1420940CDI  
US-09-974-298-184

|                       |                 |   |                   |
|-----------------------|-----------------|---|-------------------|
| Query Match           | Score 3216;     | DB 9;   | Length 619;       |
| Best Local Similarity | 98.4%           | Pred. No. 2.5e-281;   |                   |
| Matches 610;          | Conservative 2; | Mismatches 2;   | Indels 6; Gaps 2; |
| Oy                    | 1               | MSGIKKQKTEKQKSTNVVYQAHVSRNRKQGVGTGRGCTVWLTGLSGAGKTTISF      | 60                |
| Db                    | 1               | MSGIKKQKTEKQKSTNVVYQAHVSRNRKQGVGTGRGCTVWLTGLSGAGKTTISF      | 60                |
| Oy                    | 61              | ALAEYLVSHTAIPCYSLDGDVNRHGLNRNLGSPGDRENNIRIAEVAKLFDAGLVCITS  | 120               |
| Db                    | 61              | ALAEYLVSHTAIPCYSLDGDVNRHGLNRNLGSPGDRENNIRIAEVAKLFDAGLVCITS  | 120               |
| Oy                    | 121             | FISFPAKDRENAKIHESAGLPFFEIFVDAPLNTICESRDVKGLYKRAGEIKGFTGIDS  | 180               |
| Db                    | 121             | FISFPAKDRENAKIHESAGLPFFEIFVDAPLNTICESRDVKGLYKRAGEIKGFTGIDS  | 180               |
| Oy                    | 181             | DYKPEPTEPVRLKTNLSTVSDCVHQQVVELLOPNIPTVTI IKDHELFPVKNKLDHVAE | 240               |
| Db                    | 181             | DYKPEPTEPVRLKTNLSTVSDCVHQQVVELLOPNIPTVTI IKDHELFPVKNKLDHVAE | 240               |



Db 609 KAWTVLTYEYKS 620

RESULT 3

US-09-966-262-196  
; Sequence 196, Application US/09966262  
; Publication No. US20030050461A1

GENERAL INFORMATION:

; APPLICANT: Young et al.

; TITLE OF INVENTION: 87 Human Secreted Proteins

; FILE REFERENCE: P2004P1

; CURRENT APPLICATION NUMBER: US/09/966,262

; CURRENT FILING DATE: 2001-10-01

; PRIOR APPLICATION NUMBER: US 09/154,707

; PRIOR FILING DATE: 1998-09-17

; PRIOR APPLICATION NUMBER: PCT/US98/05311

; PRIOR FILING DATE: 1998-03-19

; PRIOR APPLICATION NUMBER: US 60/041,277

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/042,344

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,276

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,281

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/048,094

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,350

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,188

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,135

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/050,937

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,187

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,099

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,352

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,186

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,069

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,095

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,131

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,096

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,351

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/048,154

; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/054,804

; PRIOR FILING DATE: 1997-08-05

; PRIOR APPLICATION NUMBER: US 60/056,370

; PRIOR FILING DATE: 1997-08-19

; PRIOR APPLICATION NUMBER: US 60/060,862

; PRIOR FILING DATE: 1997-10-02

; NUMBER OF SEQ ID NOS: 343

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 196

; LENGTH: 624

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-966-262-196

Query Match 79.5%; Score 2602; DB 9; Length 624;  
Best Local Similarity 77.1%; Pred. No. 6.1e-226;  
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 5   | KKQKTEN-----QOKSTNVVYQAHVHSRNRKRGQVVGTRGGFRGCTVWLTGLSGAGKTTI  | 58  |
| Db | 9   | KKVKLSNNAQNMGMORATNTVYQAHVHSRNRKRGQVVGTRGGFRGCTVWLTGLSGAGKTTV | 68  |
| Qy | 59  | SFALEEVLVSHAIPCYSLDGNVRHGLNRNLGFSFGDREENIRRIAEYAKIFADAGLVCI   | 118 |
| Db | 69  | SMALEEVLVCHGIFCYTLDDGNIRGLNKNLGLFSFDEENVRRIAEYAKIFADAGLVCI    | 128 |
| Qy | 119 | TSFTSPFAKORENARKIHESAGLPFFETFDVADUNICESRDVKGLYKARAGEIKGFTGI   | 178 |
| Db | 129 | TSFISPTQDRNNARQIHGASLPFFETFDVADUNICESRDVKGLYKARAGEIKGFTGI     | 188 |
| Qy | 179 | DSYKEKPEPVLKTNLSTVSDCVHVVQVVELLQEQNIPTIIDIHELFPENKLDHVR       | 238 |
| Db | 189 | DSEYKEPELVLKTDSCDVNCVQVVELLQEQNIPTIIDIHELFPENKLDHVR           | 248 |
| Qy | 239 | AEATLPSLSITKLDLQWVQVLSGQWATPLKGFMRKEYLQVHMFDTLDDGVINMSIPI     | 298 |
| Db | 249 | TAETLPALKINKVDMQVQVLAEGWATPLNGFMRKEYLQVHMFDTLDDGVINLSYPI      | 308 |
| Qy | 299 | VLPSVADKTRLEGCSEFVLAHGGRRVAILRDADFEYEHKRECSRVWGTCTTKPHIKM     | 358 |
| Db | 309 | VLTAHEDKERLDGCTAFALMYEGRRVAILRNPEFFEHKRECSRVWGTCTTKPHIKM      | 368 |
| Qy | 359 | VMSGDMLVGGDLQVLEKIRWNGDLOVRLTPELKKCKEMNADAVAFOLRNPVHNGH       | 418 |
| Db | 369 | VMEQDMLIGDGLQVLRVWNGDLOVRLTPELKKCKEMNADAVAFOLRNPVHNGH         | 428 |
| Qy | 419 | ALLMQDTCRRLLERGYKHPVLLHPLGQWTKDDVPLDWRMKQHAALVEEGVLDPKSTIV    | 478 |
| Db | 429 | ALLMQDTHKQLLERYGRRPVLLHPLGQWTKDDVPLDWRMKQHAALVEEGVLDPKSTIV    | 488 |
| Qy | 479 | AIFSPMLYAGPTEVQWCHRCRMTAGANFYIVGRDPAGMPHETGKDIYEPHSGKVLSM     | 538 |
| Db | 489 | AIFSPMLYAGPTEVQWCHRCRMTAGANFYIVGRDPAGMPHETGKDIYEPHSGKVLSM     | 548 |
| Qy | 539 | APGLTSVEIIPFRVAAYNAKAKAMDYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP    | 598 |
| Db | 549 | APGLTLEIVFRVAAYNAKAKAMDYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP      | 608 |
| Qy | 599 | KAWKVLTDYYS 610   |     |
| Db | 609 | KAWTVLTYEYKS 620  |     |

RESULT 4

US-09-983-966-196  
; Sequence 196, Application US/09983966  
; Publication No. US20030060619A1

GENERAL INFORMATION:

; APPLICANT: Young et al.

; TITLE OF INVENTION: 87 Human Secreted Proteins

; FILE REFERENCE: P2004P1

; CURRENT APPLICATION NUMBER: US/09/983,966

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 09/154,707

; PRIOR FILING DATE: 1998-09-17

; PRIOR APPLICATION NUMBER: PCT/US98/05311

; PRIOR FILING DATE: 1998-03-19

; PRIOR APPLICATION NUMBER: US 60/041,277

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/042,344

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,276

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/041,281

; PRIOR FILING DATE: 1997-03-21

; PRIOR APPLICATION NUMBER: US 60/048,094

; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: US 60/048,350  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,188  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,135  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/050,937  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,187  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,099  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,352  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,186  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,069  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,095  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,131  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,096  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,355  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,160  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,351  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/048,154  
;; PRIOR FILING DATE: 1997-05-30  
;; PRIOR APPLICATION NUMBER: US 60/054,804  
;; PRIOR FILING DATE: 1997-08-05  
;; PRIOR APPLICATION NUMBER: US 60/056,370  
;; PRIOR FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: US 60/060,862  
;; PRIOR FILING DATE: 1997-10-02  
;; NUMBER OF SEQ ID NOS: 343  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 196  
;; LENGTH: 624  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-898-966-196

Query Match 79.5%; Score 2602; DB 9; Length 624;  
Best Local Similarity 77.1%; Pred. No. 6.1e-226;  
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;  
QY 5 KKOKTEN-----QOKSTNVVYQAAHVSRNKRQGVVGTGRGGFRCGCTVWLTGLSGAGKTTI 58  
DB 9 KVKLSNNAQNMGMORATNVTYQAAHVSRNKRQGVVGTGRGGFRCGCTVWLTGLSGAGKTTV 68  
QY 59 SFALPEYLVSHPICVSLDGNVHGLNRLNGLFSPGDRREENRRIAEVAKLFADAGLVCI 118  
DB 69 SWALEEYLCHGIPCYTLDGDNIRQGLNKLNGFSPEDREENVRRIAEVAKLFADAGLVCI 128  
QY 119 TSFISPFADRENARKIHESAGLPFEIEVDAPLNICESRDVKGKLYKARAGEIKGFTGI 178  
DB 129 TSFISPYTDORNARQIHESAGLSIPFEIEVDAPLVHCEQDVQVYVLLQERDIPVDASYEYKELYVPENKHLAK 188  
QY 179 DSDYEKPEPPELVKTNLSTVSDCVHQVVELLQEQNIVPTTIKDIHELFPVENKLDHVR 238  
DB 189 DSEYEKPEAPELVKLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248  
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEKEYLQVMHFTTLLDDGVINNSIPI 298  
DB 189 DSEYEKPEAPELVKLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248  
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEKEYLQVMHFTTLLDDGVINNSIPI 298  
DB 249 TDAETLPALKINKVDQMOWVQVLAEGWATPLNGFMREYRLQCLHFDCLDDGGVNLSPVI 308  
QY 299 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 358  
DB 299 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 358  
QY 358 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 358  
DB 309 VLTATHEDKERLDGGCTAFALMYEGRVAILRNPFEFHEKREKERCARQWGTTCCKNHPYIKM 418  
DB 368 VMEQDGLWIGDGLQVLDVRYWNGDGLQYRLTPELQKQKFKDMNADAVFAFQLRNPVHNGH 428

QY 359 VMESGDWLVGGDLQVLEKIRWINDGLQYRLTPELQKQKCKEMNADAVFAFQLRNPVHNGH 418  
DB 369 VMEQDGLWIGDGLQVLDVRYWNGDGLQYRLTPELQKQKFKDMNADAVFAFQLRNPVHNGH 428  
QY 419 ALLMQDTCRRLLERGERYKHPVLLHPLGGWTKDDVPLDMRMKQHAHAALVEGVLDPKSTIV 478  
DB 429 ALLMQDTHKOLLERGERYRPPVLLHPLGGWTKDDVPLDMRMKQHAHAALVEGVLPETTVV 488  
QY 479 AIFPSPMLYAGTEVOMHCHRSRMIAAGANEYIVGRDPAGMPHPETKKDLYEPHGGVLSM 538  
DB 489 AIFPSPMYAGTEVOMHCHRCARMVAGANFYIVGRDPAGMPHPETGKDLPEPSHGAKVLTM 548  
QY 539 APGLTSEIIPFRAVAYNKAAMDFYDPAHNEFDFISGTRMRKRLAREGENPPDGFMAP 598  
DB 549 APGLITLPIVPRVAAYNKKKRMDDYDSEHHEDEFEISGTRMRKRLAREGCKPPEGFMAP 608  
QY 599 KAWKVLTDIYRS 610  
DB 609 KAWTVLTYEYKS 620

RESULT 5  
US-10-143-090-196  
; Sequence 196, Application US/10143090  
; Publication No. US20030069406A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004PI  
; CURRENT APPLICATION NUMBER: US/10/143,090  
; PRIOR FILING DATE: 2002-05-13  
; PRIOR APPLICATION NUMBER: 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 196  
; LENGTH: 624  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-143-090-196

Query Match 79.5%; Score 2602; DB 9; Length 624;  
Best Local Similarity 77.1%; Pred. No. 6.1e-226;  
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QOKSTNVVYQAAHVSRNKRQGVVGTGRGGFRCGCTVWLTGLSGAGKTTI 58  
DB 9 KVKLSNNAQNMGMORATNVTYQAAHVSRNKRQGVVGTGRGGFRCGCTVWLTGLSGAGKTTV 68  
QY 59 SFALPEYLVSHPICVSLDGNVHGLNRLNGLFSPGDRREENRRIAEVAKLFADAGLVCI 118  
DB 69 SWALEEYLCHGIPCYTLDGDNIRQGLNKLNGFSPEDREENVRRIAEVAKLFADAGLVCI 128  
QY 119 TSFISPFADRENARKIHESAGLPFEIEVDAPLNICESRDVKGKLYKARAGEIKGFTGI 178  
DB 129 TSFISPYTDORNARQIHESAGLSIPFEIEVDAPLVHCEQDVQVYVLLQERDIPVDASYEYKELYVPENKHLAK 188  
QY 179 DSDYEKPEPPELVKTNLSTVSDCVHQVVELLQEQNIVPTTIKDIHELFPVENKLDHVR 238  
DB 189 DSEYEKPEAPELVKLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248  
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEKEYLQVMHFTTLLDDGVINNSIPI 298  
DB 249 TDAETLPALKINKVDQMOWVQVLAEGWATPLNGFMREYRLQCLHFDCLDDGGVNLSPVI 308  
QY 299 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 358  
DB 299 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 358  
QY 358 VLPVSAEDKTRLEGCKSFVLAHGGRRVAILRDAEFYEHKREKERSRVWGTTCTCKPHIKM 418  
DB 368 VMEQDGLWIGDGLQVLDVRYWNGDGLQYRLTPELQKQKFKDMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRRLLERKYPHVPVLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 478  
Db 429 ALLMODTHKQLLERGYRRPVLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 488  
QY 479 AIFPSPMLYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHETPKDLYEPHGGKVLMS 538  
Db 489 AIFPSPMYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHETPKDLYEPHGGKVLMS 548  
QY 539 APLGTSVELIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 598  
Db 549 APLGTLLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 608  
QY 599 KAWKVLTDYRS 610  
Db 609 KAWTVLTYEYS 620

RESULT 6  
US-09-771-161A-244  
; Sequence 244, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771,161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 133619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 244  
; LENGTH: 624  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-244

Query Match 79.1%; Score 2588; DB 10; Length 624;  
Best Local Similarity 76.6%; Pred. No. 1.1e-224;  
Matches 469; Conservative 68; Mismatches 69; Indels 6; Gaps 1;  
QY 5 KKQKTEN-----QOKSTNVYQAHVSRNKRQGVVGRGGRGCTVMTGLSGAGKTTI 58  
Db 9 KVKLSNNAQNWGMQRATNVTYQAHVSRNKRQGVVGRGGRGCTVMTGLSGAGKTTI 68  
QY 59 SPALAEYLVSHPICYSIDGDNVNRHGLNRNLGSPGDRNTRRTAEVAKLFADAGLVCI 118  
Db 69 SMALEYLVCHPICYSIDGDNVNRHGLNRNLGSPGDRNTRRTAEVAKLFADAGLVCI 128  
QY 119 TSFISPEAKDRNARKIHESAGLPFFELFVDAPLNICESRDYKGLYKRAAGEIKGFTCI 178  
Db 129 TSFISPYQDRNNAKQIHESAGLPFFELFVDAPLNICESRDYKGLYKRAAGEIKGFTCI 188  
QY 179 DSDYKPEPTEPVLKTNLSTSDCVHQVVELLQEQNIPTTIKDIHELFPENKLDHVR 238  
Db 189 DSEYKPEAPELVKLTSDCVNDVCQVVELLQEQNIPTTIKDIHELFPENKLDHVR 248  
QY 239 AEAETPLSLITKLDQVQLVSEGWATPLKGFMEKEYLVQVHFTDLDGCVINMSPTI 298  
Db 249 TDAETPLKINKVQVQVLAEGWATPLKGFMEKEYLVQVHFTDLDGCVINMSPTI 308  
QY 299 VLPVSAEDKTRLEGCSKFLAHGGRVAILRDAAEYHKEKRCRSRVGCTTCKPHLKM 358  
Db 309 VUTATHEKRLDGGCTAFLMTEGRRVAILRNPEFEHKEKRCARQWGTCTCKNHPYIKM 368  
QY 359 VMEQGDWLVGDLQVLEIRWNGDLDQYRLTPELQKQCKEMNADAVAFQLRNPNVHGH 418  
Db 369 VMEQGDWLVGDLQVLEIRWNGDLDQYRLTPELQKQCKEMNADAVAFQLRNPNVHGH 428

QY 419 ALLMODTCRRLLERKYPHVPVLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 478  
Db 429 ALLMODTHKQLLERGYRRPVLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 488  
QY 479 AIFPSPMLYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHETPKDLYEPHGGKVLMS 538  
Db 489 AIFPSPMYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHETPKDLYEPHGGKVLMS 548  
QY 539 APLGTSVELIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 598  
Db 549 APLGTLLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 608  
QY 599 KAWKVLTDYRS 610  
Db 609 KAWTVLTYEYS 620

RESULT 7  
US-10-156-239-40  
; Sequence 40, Application US/10156239  
; Publication No. US20030036074A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Tr  
; TITLE OF INVENTION: Appase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A  
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor  
; FILE REFERENCE: 35800/247645  
; CURRENT APPLICATION NUMBER: US/10/156,239  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 09/7795,693  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/185,906  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: 09/809,557  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/192,018  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 09/808,568  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/191,790  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 09/808,767  
; PRIOR FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/191,781  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pfam consensus sequence  
US-10-156-239-40

Query Match 17.5%; Score 571.5; DB 9; Length 162;  
Best Local Similarity 67.9%; Pred. No. 1.1e-43;  
Matches 110; Conservative 17; Mismatches 32; Indels 3; Gaps 1;  
QY 41 RCTVMTGLSCAGKTTISFALAEYLVSHPICYSIDGDNVNRHGLNRNLGSPGDRNTR 100  
Db 1 RCTVMTGLSCAGKTTISFALAEYLVSHPICYSIDGDNVNRHGLNRNLGSPGDRNTR 60  
QY 101 RRIAEVAKLFADAGLVCIISFISPAKDRNARKTH---ESAGLPFFELFVDAPLNICES 157  
Db 61 RRVGEVAKLFADAGLVCIISFISPAKDRNARKTH---ESAGLPFFELFVDAPLNICES 120  
QY 158 RRVKGLYKRAAGEIKGFTGIDSDYKPEPTEPVLKTNLSTV 199  
Db 121 RDPKGLYKRAAGEIKGFTGIDSDYKPEPTEPVLKTNLSTV 162

RESULT 8  
US-10-199-485-40  
; Sequence 40, Application US/10199485  
; Publication No. US20030077626A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; APPLICANT: Silos-Santiago, Inmaculada  
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and  
; TITLE OF INVENTION: 32613, No. US20030077626A1el Human Transporters  
; FILE REFERENCE: 35800/249468  
; CURRENT APPLICATION NUMBER: US/10/199,485  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 09/795,693  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/185,906  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pfam consensus sequence  
US-10-199-485-40

Query Match 17.5%; Score 571.5; DB 9; Length 162;  
Best Local Similarity 67.9%; Pred. No. 1.1e-43;  
Matches 110; Conservative 17; Mismatches 32; Indels 3; Gaps 1;  
QY 41 RGTAVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 157  
DB 1 RGTAVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 60  
QY 101 RRTAEVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 157  
DB 61 RRVGEVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 157  
QY 158 RDVKGlyKARAGEIKGFTGIDSDEYKPTPERVLKTNLSTV 199  
DB 121 RDPKGLYKARAGEIKGFTGIDSDEYKPTPERVLKTNLSTV 162

RESULT 9  
US-09-795-693-40  
; Sequence 40, Application US/09795693  
; Patent No. US20020068710A1  
; GENERAL INFORMATION:  
; APPLICANT: Glucksmann, Maria A.  
; TITLE OF INVENTION: 20685, 579, 17114, 23821, 33894, and  
; TITLE OF INVENTION: 32613, No. US20020068710A1el Human Transporters  
; FILE REFERENCE: 35800/209292  
; CURRENT APPLICATION NUMBER: US/09/795,693  
; CURRENT FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/185,906  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Pfam consensus sequence  
US-09-795-693-40

Query Match 17.5%; Score 571.5; DB 10; Length 162;  
Best Local Similarity 67.9%; Pred. No. 1.1e-43;  
Matches 110; Conservative 17; Mismatches 32; Indels 3; Gaps 1;  
QY 41 RGTAVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 157

DB 1 RGTAVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 60  
QY 101 RRTAEVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 157  
DB 61 RRVGEVAKLADAGLVCTISFISPAKADRENARKIH---ESAGLPFFEIFVDAPLNICES 120  
QY 158 RDVKGlyKARAGEIKGFTGIDSDEYKPTPERVLKTNLSTV 199  
DB 121 RDPKGLYKARAGEIKGFTGIDSDEYKPTPERVLKTNLSTV 162

RESULT 10  
US-09-741-669-457  
; Sequence 457, Application US/09741669  
; Patent No. US20020022718A1  
; GENERAL INFORMATION:  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; TITLE OF INVENTION: Genes identified as required for  
; TITLE OF INVENTION: proliferation of E. coli  
; FILE REFERENCE: ELITRA.009A  
; CURRENT APPLICATION NUMBER: US/09/741,669  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 60/173005  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 457  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-741-669-457

Query Match 14.6%; Score 476.5; DB 10; Length 201;  
Best Local Similarity 50.2%; Pred. No. 5.6e-35;  
Matches 101; Conservative 35; Mismatches 56; Indels 9; Gaps 4;  
QY 17 NVYQAHVSRNRKQGVGTGRGCTVMTGLSGAGKTTISFALDEYLVSHAIPCYSYL 76  
DB 7 NVVHSHPTVQQR---ELHHGRGVLVWFTGLSGGKSTVAGALEALHKLGVSTYLL 62  
QY 77 DGDVNRHGLNRLGFSFGDRENRIRAEVAKLADAGLVCTISFISPAKADRENAR-KI 135  
DB 63 DGDVNRHGLNRLGFSFGDRENRIRAEVAKLADAGLVCTISFISPAKADRENAR-KI 122  
QY 136 HESAGLPFFEIFVDAPLNICESRDVKGlyKARAGEIKGFTGIDSDEYKPTPERVLKTN 195  
DB 123 GEGR---FIEVFDVDTPLAICEARDPKGLYKARAGELRNFTGIDSVEAPESAEIHLNGE 179  
QY 196 LSTVSDCVHQVVELLOEQNIV 216  
DB 180 -QLVTNLVQQLDLLLRQNDII 199

RESULT 11  
US-09-864-761-39981  
; Sequence 39981, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26

```
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39981
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006191.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: O95340, EVALUE 4.00e-33
; OTHER INFORMATION: EST_HUMAN HIT: BE292722.1, EVALUE 4.00e-32
US-09-864-761-39981
```

```
Query Match 10.4%; Score 340; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 EEYLVSHPAICYSLDGDNVRHGLNRLNFGSPGDRNIRRTAEVAKLFADAGLVCITISFI 122
Db 1 EEYLVSHPAICYSLDGDNVRHGLNRLNFGSPGDRNIRRTAEVAKLFADAGLVCITISFI 60

Qy 123 SPFAK 127
Db 61 SPFAK 65
```

```
RESULT 12
US-09-809-391-635
; Sequence 635, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
```

```
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-635
```

```
Query Match 9.9%; Score 323; DB 9; Length 174;
Best Local Similarity 83.3%; Pred. No. 3.1e-11;
Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 MSGIKKQKTEQKSTNNVYQAHVSRNKRQGVYVTRGCGTCTVWLTGLSGAGTITISF 60
Db 55 MSGIKKQKTEQKSTNNVYQAHVSRNKRQGVYVTRGCGTCTVWLTGLSGAGKNDKF 114

Qy 61 ALEEYLVSHAIPCYSLDG 78
Db 115 CPGGVLYSHAIPVNSWMG 132
```

```
RESULT 13
US-09-867-550-96
; Sequence 96, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-96
```

```
Query Match 6.8%; Score 221.5; DB 10; Length 163;
Best Local Similarity 41.9%; Pred. No. 4.1e-12;
Matches 57; Conservative 17; Mismatches 59; Indels 3; Gaps 2;

Qy 41 RGCITVWLTGLSGAGKTTISFALEEVLSHAIPCYSLDGDNVRHGLNRLNFGSPGDRN 99
Db 26 QGFTILLSGLSNCKDQIASALQVKLHEQGSRVSLLLGDTVYRHSLSSELGFSPEDRHN 85

Qy 100 IRTIAEVAKLFPADAGLVCITISFIPFAKDRNARK--THESAGLPFFEIFVDAPLNICES 157
Db 86 LQRTAFVASELSTRAGAASPIAPYEHRSQMRDITLKGGSNNFFLIHVATPLEYCEK 145

Qy 158 RDVKGLYKRAGEIK 173
Db 146 HRRRGNYAKARRGEVK 161
```

```
RESULT 14
US-09-984-245-340
; Sequence 340, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
```





Db :|||||||||||||:||||||||||||||||||||||||||||| |||  
1 MYAGTEVQWCHCRAMWAGANFYIVGRDPAGMHPHETGKDL 41

Search completed: May 30, 2003, 12:28:53  
Job time : 49 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:15:39 ; Search time 43 Seconds  
(without alignments)  
1374.946 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSGIKKQKTKNQKSTNVV.....MAPKAMKVLTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73.\*

1: pirl.\*

2: pirl2.\*

3: pirl3.\*

4: pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description         |
|------------|--------|-------------|--------|----------|---------------------|
| 1          | 2602   | 79.5        | 624    | 1 JW0087 | adenyl-yl-sulfate k |
| 2          | 2367   | 72.4        | 610    | 1 JC4383 | adenyl-yl-sulfate k |
| 3          | 1887   | 57.7        | 652    | 1 T24918 | 3'-phosphoadenosin  |
| 4          | 1299   | 39.7        | 424    | 2 S44079 | sulfate adenyl-yltr |
| 5          | 1282   | 39.2        | 465    | 2 E71409 | sulfate adenyl-yltr |
| 6          | 1274   | 38.9        | 463    | 2 S44267 | sulfate adenyl-yltr |
| 7          | 1268.5 | 38.8        | 476    | 2 S44943 | sulfate adenyl-yltr |
| 8          | 1265   | 38.7        | 483    | 2 T14475 | sulfate adenyl-yltr |
| 9          | 1259.5 | 38.5        | 469    | 2 T52659 | sulfate adenyl-yltr |
| 10         | 1230.5 | 37.6        | 489    | 1 T01204 | sulfate adenyl-yltr |
| 11         | 1212.5 | 37.1        | 459    | 2 T08594 | probable sulfate a  |
| 12         | 556    | 17.0        | 202    | 2 T50101 | adenyl-yl-sulfate k |
| 13         | 551.5  | 16.9        | 200    | 2 E96912 | adenyl-yl-sulfate k |
| 14         | 523    | 16.0        | 202    | 1 S17244 | adenyl-yl-sulfate k |
| 15         | 500    | 15.3        | 199    | 2 A69839 | adenyl-yl-sulfate k |
| 16         | 497.5  | 15.2        | 293    | 2 T06100 | adenyl-yl-sulfate k |
| 17         | 493.5  | 15.1        | 213    | 2 A50408 | adenyl-yl-sulfate k |
| 18         | 487.5  | 14.9        | 202    | 2 A83836 | adenyl-yl-sulfate k |
| 19         | 481.5  | 14.7        | 202    | 2 A10856 | adenosine 5'-phosp  |
| 20         | 481.5  | 14.7        | 635    | 2 A87433 | hypothetical prote  |
| 21         | 476.5  | 14.6        | 201    | 1 B65056 | adenyl-yl-sulfate k |
| 22         | 476.5  | 14.6        | 201    | 2 D51079 | adenosine 5'-phosp  |
| 23         | 476.5  | 14.6        | 201    | 2 E85924 | adenosine 5'-phosp  |
| 24         | 476    | 14.6        | 276    | 1 S47640 | adenyl-yl-sulfate k |
| 25         | 469.5  | 14.4        | 312    | 2 T08076 | adenyl-yl-sulfate k |
| 26         | 463    | 14.2        | 215    | 2 F82062 | adenyl-yl-sulfate k |
| 27         | 456.5  | 14.0        | 614    | 1 B70772 | probable adenyl-yl  |
| 28         | 451.5  | 13.8        | 197    | 2 C69877 | adenyl-yl-sulfate k |
| 29         | 440    | 13.5        | 644    | 2 AD3471 | adenyl-yl-sulfate k |

#### ALIGNMENTS

##### RESULT 1

JW0087

adenyl-yl-sulfate kinase (EC 2.7.1.25) - human

N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS

N:Contains: adenyl-yl-sulfate kinase (EC 2.7.1.25); sulfate adenyl-yltransferase (EC 2.7

C:Species: Homo sapiens (man)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 03-Jun-2002

C:Accession: JW0087

R:Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakayama, T.; Nakajima, H.

Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998

A:Title: cDNA cloning, expression, and characterization of the human bifunctional ATP

A:Reference number: JW0087; MUID:98312048; PMID:9648242

A:Accession: JW0087

A:Molecule type: mRNA

A:Residues: 1-624 <Y>

A:Cross-references: GB:AR033026; NID:G3378100; PIDN:AAC28429.1; PID:G3378101

A:Experimental source: Brain

C:Function: <ASKF>

A:Description: as adenyl-yl-sulfate kinase catalyzes the phosphorylation of adenyl-ylsal

C:Function: <SATE>

A:Description: as sulfate adenyl-yltransferase catalyzes the reaction of sulfate and A

C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenyl-ylsulfa

C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop

F:52-215/Domain: adenyl-yl-sulfate kinase homology <ASK>

F:59-66/Region: nucleotide-binding motif A (P-loop)

F:226-620/Domain: sulfate adenyl-yltransferase homology <SAT>

F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 79.5% Score 2602; DB 1; Length 624;

Best Local Similarity 77.1%; Pred. No. 6.4e-186;

Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKQTKEN-----QOKSTNVVYQAHVSRNKRGOVVGTRGGPRGCTVWLTGLSGAGKTTI 58

DB 9 KKLKLSNAQNWGMQRATNVITQAHVSRNKRGOVVGTRGGPRGCTVWLTGLSGAGKTTV 68

QY 59 SFALFEYLVSHAIPCYSLDGDGNVRHGLNRNLGFSFGDREENTRIRAEVAKLFADAGLVCI 118

DB 69 SMALEEYLVCHGICPYTLTDGDNIRGLNKLGLFSFGDREENTRIRAEVAKLFADAGLVCI 128

QY 119 TSFTSPFAKDRENARKIHESAGLPFEIFVDAPLNCESRDVKGILYKARAGEIKGFTGI 178

DB 129 TSFISPYTQDRNRRQIHEGASLPFEFVVDAPLHVCQRDQVKGILYKARAGEIKGFTGI 188

QY 179 DSDVEKPTPERVLKTNLSTVSDCVHVVQLQEQNIYPYTIKDIHELFPENKLDHVR 238

DB 189 DSEYKPEAPLEKLTDSKDNDCVQVVELLQERDIPVDVASYEVKELYPENKHLAK 248

QY 239 AEATFLPSLSTTKLDLQWQVLSQGWATPLKGFMEKEYLVQVMHFDITLDDGVINNSIPTI 298

DB 249 TDAETLPALKINKVDNQWQVLAEGWATPLNGFMREYVQLQCLHFDCLLDGGVINLSVPI 308

probable adenyl-yl-  
adenyl-yl-sulfate k  
adenyl-yl-sulfate k  
adenyl-yl-sulfate k  
adenyl-yl-sulfate k  
ATP sulfurylase, l  
probable adenyl-yl-  
adenyl-yl-sulfate k  
adenosine 5'-phosp  
sulfate adenyl-yltr  
adenyl-yl-sulfate k  
sulfate adenyl-yltr  
sulfate adenyl-yltr  
sulfate adenyl-yltr  
sulfate adenyl-yltr

30 438 13.4 633 2 H95932  
31 436 13.3 641 1 ZZRNQ  
32 436 13.3 641 2 E95320  
33 434.5 13.3 206 2 H84978  
34 429 13.1 208 2 A84073  
35 425 13.0 660 2 G82672  
36 418 12.8 546 1 C70393  
37 415 12.7 620 1 I39755  
38 407 12.4 196 2 H83472  
39 400.5 12.2 387 2 D75594  
40 396 12.1 177 1 S74917  
41 388 11.9 381 2 G83835  
42 384 11.7 214 1 S18729  
43 374 11.4 574 1 S55034  
44 370.5 11.3 379 2 B84073  
45 368 11.3 573 1 A53651

QY 299 VLPVSAEDKTRLEGGCKSFVLAHGRVVAIRDAEFAEYEHKREKCSRVMGTTCTKHPHMK 358  
Db 309 VLTATHDEKRLDGCATFALMYGERRVAIRLNPEFFEHKREKCAROWGTTCKNHPYIKM 368  
QY 359 VMESGDMVLGGDLQVLEKIRWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVNHG 418  
Db 369 VMQGDWLLGGDLQVLDVYNDGDLQYRUTPELKOKFKDMNADAVAFQLRNPVNHG 428  
QY 419 ALLMODTCRLLRGYKHPVLLHPLGGWTKDDVPLDWRMKQHAALVEGVLDPKSTIV 478  
Db 429 ALLMODTHKLLRGYRPPVLLHPLGGWTKDDVPLDWRMKQHAALVEGVLDPKSTIV 488  
QY 479 AIFPSPMLYAGTPEVQVHCRSMIAGANFYIVGRDPAGMPHPTETKDLIYPTGGKVLMS 538  
Db 489 AIFPSPMLYAGTPEVQVHCRSMIAGANFYIVGRDPAGMPHPTETKDLIYPTGGKVLMS 548  
QY 539 APLGTSVEIIPFVAAYNAKAMDFYDPAHNEFDYISGTRMKRLAREGENPDGFMAP 598  
Db 549 APLGLTLEIVPFRVAAYNAKAMDFYDPAHNEFDYISGTRMKRLAREGENPDGFMAP 608  
QY 599 KAWKVLTDYYS 610  
Db 609 KAWTVLTEYYS 620

RESULT 2  
JC4383  
adenyl-yl-sulfate kinase (EC 2.7.1.25) - spoonworm (Urechis caupo)  
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS  
N:Contains: adenyl-yl-sulfate kinase (EC 2.7.1.25); sulfate adenyl-yl-transferase (EC 2.7.7.4)  
C:Species: Urechis caupo  
C:Date: 17-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 03-Jun-2002  
C:Accession: JC4383  
R:Rosenthal, E.; Leustek, T.  
Gene 165, 243-248, 1995  
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfurylase and adenylyl-sulfate kinase activities  
A:Reference number: JC4383; MUID:96096529; PMID:8522184  
A:Accession: JC4383  
A:Molecule type: mRNA  
A:Residues: 1-610 <ROS>  
A:Cross-references: GB:I39001; NID:g705384; PIDN:AAB00139.1; PID:g705385  
C:Genetics:  
A:Gene: paps  
C:Function: <ASAP>  
A:Description: as sulfate adenyl-yl-sulfate kinase catalyzes the phosphorylation of adenyl-yl-sulfate  
C:Function: <SATF>  
A:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenyl-yl-sulfate  
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;  
F:37-200/Domain: adenyl-yl-sulfate kinase homology <ASK>  
F:44-51/Region: nucleotide-binding motif A (P-loop)  
F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PAB>  
F:211-605/Domain: sulfate adenyl-yl-transferase homology <SAT>  
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 72.4%; Score 2367; DB 1; Length 610;  
Best Local Similarity 71.4%; Pred. No. 2.1e-168;  
Matches 432; Conservative 69; Mismatches 104; Indels 0; Gaps 0;

QY 11 NOCKSTNVVYQAHVSRNRKQGVGTRGGGCTVMTGLSGAGKTTISFALEEYLVSHA 70  
Db 6 NGOLATNVTFQTHVSRNRKQGVGTRGGGCTVMTGLSGAGKTTISFALEEYLVSG 65

QY 71 IPCYSLDGNVRHGLNRNLGFSQDREENIRIAEVAKLFDAGLVLCITISFPAKRE 130  
Db 66 IPTYSLDGNVRHGLNRNLGFSQDREENIRIAEVAKLFDAGLVLCITISFPAKRE 125

QY 131 NARKIHESAGLPPFEIFVDAPLNCESRDVKGILYKARAGEIKGTFIDSDYKPTPER 190  
Db 126 LARSLHEQAGLPPFEIFVDAPLNCESRDVKGILYKARAGEIKGTFIDSDYKPTPER 185

QY 191 VLKTNLSVSDCVHQVVELLQEQNIYPTTIKIDIELFYPENKLDHVAEAEFLPSLIT 250  
Db 185 VLKTNLSVSDCVHQVVELLQEQNIYPTTIKIDIELFYPENKLDHVAEAEFLPSLIT 245

Db 186 QLYAGNKSIDECVQEVVSVLLQKNGVWPESAVNIVKELFVPESGLEHAKAEIVDLPMTMEIT 245  
QY 251 KLDLOWQVLSSEGATPLKGFMEKEKEYLVQVHFDFTLLDDGVINNSIPVLPVSAEDKTRL 310  
Db 246 KLDLOWQVLSSEGATPLKGFMEKEKEYLVQVHFDFTLLDDGVINNSIPVLPVSAEDKTRL 305  
QY 311 EGCSKEVLAHGGRRVAIRDAEFAEYEHKREKCSRVMGTTCTKHPHMKMWSGDLVGGD 370  
Db 306 EGCSKEVLAHGGRRVAIRDAEFAEYEHKREKCSRVMGTTCTKHPHMKMWSGDLVGGD 365  
QY 371 LQVLEKIRWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVNHGHALMODTCRLL 430  
Db 366 LEVLEIRITWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVNHGHALMODTCRLL 425  
QY 431 ERGYKHPVLLHPLGGWTKDDVPLDWRMKQHAALVEGVLDPKSTIVAFSPMLYAGP 490  
Db 426 ERGYKHPVLLHPLGGWTKDDVPLDWRMKQHAALVEGVLDPKSTIVAFSPMLYAGP 485  
QY 491 TEVQVHCRSMIAGANFYIVGRDPAGMPHPTETKDLIYPTGGKVLMSAPGLTSVEIIP 550  
Db 486 TEVQVHCRSMIAGANFYIVGRDPAGMPHPTETKDLIYPTGGKVLMSAPGLTSVEIIP 545  
QY 551 RVAAYNKKAMDFYDPAHNEFDYISGTRMKRLAREGENPDGFMAPKAWKVLTDYYS 610  
Db 546 RVAAYNKKAMDFYDPAHNEFDYISGTRMKRLAREGENPDGFMAPKAWKVLTDYYS 605  
QY 611 EMDKN 615  
Db 606 KAQOS 610

RESULT 3  
T24918  
3'-phosphoadenosine-5'-phosphosulfate synthetase - Caenorhabditis elegans  
N:Alternate names: protein T14G10.1  
C:Species: Caenorhabditis elegans  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 19-Jan-2001  
C:Accession: T24918  
R:Wild, A.  
Submitted to the EMBL Data Library, January 1996  
A:Reference number: Z19954  
A:Accession: T24918  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-652 <WIL>  
A:Cross-references: EMBL:Z68880; PIDN:CAA93098.1; GSPDB:GN00022; CESP:T14G10.1  
A:Experimental source: clone T14G10  
C:Genetics:  
A:Gene: CESP:T14G10.1  
A:Map position: 4  
A:Introns: 23/3; 82/3; 176/1; 535/2; 623/2  
C:Function: <ASKP>  
A:Description: as adenyl-yl-sulfate kinase catalyzes the phosphorylation of adenyl-yl-sulfate  
C:Function: <SATF>  
A:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenyl-yl-sulfate  
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;  
F:36-219/Domain: adenyl-yl-sulfate kinase homology <ASK>  
F:56-70/Region: nucleotide-binding motif A (P-loop)  
F:180-190/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PA>  
F:230-647/Domain: sulfate adenyl-yl-transferase homology <SAT>  
F:137/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 57.7%; Score 1887; DB 1; Length 652;  
Best Local Similarity 56.5%; Pred. No. 1.4e-132;  
Matches 362; Conservative 100; Mismatches 151; Indels 28; Gaps 6;

QY 1 MSGIKKQKTEK-QQKSTNVVYQAHVSRNRKQGVGTRGGGCTVMTGLSGAGKTTIS 59  
Db 14 MPMKLRPVSSYSSGGSTNTITYEHTLSREERAAVHGFEGFCTIWTGLSGAGKTTIS 73

QY 60 FALEEYLVSHAIPCYSLDGNVRHGLNRNLGFSQDREENIRIAEVAKLFDAGLVLCIT 119  
Db 59 FALEEYLVSHAIPCYSLDGNVRHGLNRNLGFSQDREENIRIAEVAKLFDAGLVLCIT 114

Db 74 FALERTLANKGIPCYGLDGNIRHGLCKNLGFSKEDRQNRIRRAEVAKLFADSGMICIA 133

QY 120 SFISPAKRENAKIHESAGLPFPEFVDAPLNICESRDVGLYKRARAGETKGTGID 179

Db 134 AFISPOEDRLARKIHESNVKFEVHVSTTLEVECEQRPQLYKARAGOLGFTGID 193

QY 180 SYEKPEPTEPVLKNTLSVDCVHQVVELLQEQNIVPYTI--IKDIHFLVPEP-KLDH 236

Db 194 SAYEPPENAEIIIDAGKQGVQCVQKVLHESKGLLPEQIPDPVPAVRELFVSDDLTVAE 253

QY 237 VRAEAETPLSLTKLDLQWQVLSGFWATPLKGFMRKEKEYLQVHFDFTLLD----- 288

Db 254 LKESONLPTVELTKVDLQWLOVLAEGWATPLSGFMREROYLOSMHFGGLDLKHKVAFV 313

QY 289 -----DGVINNSIPVLPVSAEDKTRLEGCSKFVLAHGGRRVAILRDAEFY 334

Db 314 GEKSDDKEDSWPMDD-INGSPVLPISDDVKKLEGVTRALKYNGQVYALLSPEIF 372

QY 335 EHRKEBRCRVGTTCTKHPHKKMVGSDWLVGGDLQVLEKIRWNGDLDQVRLTPLELK 394

Db 373 EHRKDBRVCRQFTNDRHPVAQVLESNGLLGGDVAVVQKIQFNGDLGDKYRKTPNELR 432

QY 395 QCKENMADAVAFOLRNPHNGHALLMODTCRLLERCKYKHPVLLHPLGGLTKDDVDP 454

Db 433 AIPAENADAVAFOLRNPHNGHALLMODTCRLLERCKYKHPVLLHPLGGLTKDDVDP 491

QY 455 LDWRMKAHAALVEEGVLDPKSTVIAIFPSPMLYAGTEVQWCHRSRMIAAGANFYIVGRDP 514

Db 492 LDRIKQHEAVTAERVLDPETWVLSIFPSPMYAGTEVQWCHRSRMIAAGANFYIVGRDP 551

QY 515 AGMHPHETKDYEPHGGKVLSPMAPGLTSVEIIPFVAAYNKAAMKAMFYDPAHNEFD 574

Db 552 AGTQKPGSDALYETHTGAKVLSMAPGLSALHILPFRVAAYDKTKMGMAFFDPSRQDEE 611

QY 575 FTSCEMRKLAREGENPPDGFMAPKAWKVLTDYRSEMDKN 615

Db 612 NISGTKMRGLARNGDTPPEGEFMAPTAWEVLGYKSLQNSN 652

RESULT 4

S44079

sulfate adenylyltransferase (EC 2.7.7.4) met3-1 - potato

C:Species: Solanum tuberosum (potato)

C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999

C:Accession: S67497; S44079

R:Klonus, D.; Hoefgen, R.; Willmitzer, L.; Riesmeyer, J.W.

Plant J. 6, 105-112, 1994

A:Title: Isolation and characterization of two cDNA clones encoding ATP-sulfurylases from potato tubers

A:Reference number: S67497; MUID:95004649; PMID:7920699

A:Accession: S67497

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-424 <KLO>

A:Cross-references: EMBL:X75041; NID:g471342; PIDN:CAA52953.1; PID:g531495

C:Genetics:

A:Gene: met3-1

C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology

C:Keywords: nucleotidyltransferase

F:11-409/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 39.7%; Score 1299; DB 2; Length 424;

Best Local Similarity 61.7%; Pred. No. 5.1e-89;

Matches 240; Conservative 64; Mismatches 81; Indels 4; Gaps 4;

QY 226 ELFPENKLDHVRAREETPLSLITKLDLQWQVLSGFWATPLKGFMRKEKEYLQVHFDFT 285

Db 21 ELFVDESQDLRKRKEALNPKLTKLDVWVHVLSGFWASPLKGFMRSEYELQTLHFN 80

QY 286 L-LDDG-VTNMSTPIVLPVSAEDKTRLEGCSFVL-AHGRRVAILRDAEFYHRRKEER 342

Db 81 IRLDGSVNMSVPVLAIDDSQKNQIDGSSVALVVDGNNFALITDIEYKHKEER 140

QY 343 SRVWGTCTCKHPHI-KMVGSDWLVGGDLQVLEKIRWNGDLDQVRLTPLELKQCKEMN 401

Db 141 ARTWGTAPCLPVVDEAITHSGNWLICGDLKVIPEVKYHGDLSFRLSPSELRAEFTRRN 200

QY 402 ADAVFAFOLRNPHNGHALLMODTCRLLERCKYKHPVLLHPLGGLTKDDVPLDWRMKQ 461

Db 201 ADAVFAFOLRNPHNGHALLMODTCRLLERCKYKHPVLLHPLGGLTKDDVPLDWRMKQ 260

QY 462 HAAVLEEGVLDPKSTVIAIFPSPMLYAGTEVQWCHRSRMIAAGANFYIVGRDPAGMPHPE 521

Db 261 HEKVLEGGVLDPETTVLSIFPSPMYAGTEVQWCHRSRMIAAGANFYIVGRDPAGMPHPE 320

QY 522 TKDLYEPHGGKVLSPMAPGLTSVEIIPFVAAYNKAAMKAMFYDPAHNEFDIFISGTRM 581

Db 321 EKRDLYDADHGKVLSPMAPGLSALHILPFRVAAYDKTKMGMAFFDPSRQDEE 380

QY 582 RKLAREGENPPDGFMAPKAWKVLTDYRSEMDKN 610

Db 381 RTLAKTKEGPPDGFMAPKAWKVLTDYRSEMDKN 409

RESULT 5

E71409

sulfate adenylyltransferase (EC 2.7.7.4) precursor (clone APS3) - Arabidopsis thaliana

N:Alternate names: ATP sulfurylase

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000

C:Accession: E71409; S68201

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenger, T.; Pohl, T.M.; Terry, N.; vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; C.; Chalatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: E71409

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-465 <BEV>

A:Cross-references: GB:297336; NID:g2244788; PIDN:CAB10247.1; PID:g2244824

R:Murillo, M.; Leustek, T.

Arch. Biochem. Biophys. 323, 195-204, 1995

A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherichia coli

A:Reference number: S68024; MUID:96019964; PMID:7487067

A:Accession: S68201

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-6,'D',8-35,'A',37-465 <MUR>

A:Cross-references: EMBL:U06275; NID:g459143; PIDN:AAA92350.1; PID:g459144

C:Genetics:

A:Map position: 4COP9-4G3845

A:Gene: nuclear

C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology

C:Keywords: chloroplast; nucleotidyltransferase

F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>

F:51-465/Product: sulfate adenylyltransferase #status predicted <MAT>

F:52-450/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 39.2%; Score 1282; DB 2; Length 465;

Best Local Similarity 62.3%; Pred. No. 1.1e-87;

Matches 243; Conservative 60; Mismatches 81; Indels 6; Gaps 5;

QY 226 ELFPENKLDHVRAREETPLSLITKLDLQWQVLSGFWATPLKGFMRKEKEYLQVHFDFT 285

Db 62 DLVPEPRRREKKHEADLPVRLTAIDLQWVLSGFWASPLKGFMRSEYELQTLHFN 121

QY 286 L-LDDG-VTNMSTPIVLPVSAEDKTRLEGCSFVL-AHGRRVAILRDAEFYHRRKEER 341

Db 122 LNLDDGSVNMSPVLAIDDSQKNQIDGSSVALVVDGNNFALITDIEYKHKEER 180



## RESULT 8

T14475  
sulfate adenylyltransferase (EC 2.7.7.4) ASBo precursor - wild cabbage  
N:Alternate names: ATP sulfurylase  
C:Species: Brassica oleracea (wild cabbage)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000  
C:Accession: T14475  
R:Hatzfeld, Y.; Logan, H.M.; Cathala, N.; Davidian, J.C.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z18111  
A:Accession: T14475  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-483 <HAT>  
A:Cross-references: EMBL:U69694; NID:g1589912; PID:g1589913  
C:Genetics:  
A:Gene: ASBo  
A:Introns: 185/2; 407/3; 436/3  
C:Function:  
A:Description: catalyzes the activation of sulfate to adenylylsulfate  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: nucleotidyltransferase  
F:68-466/Domain: sulfate adenylyltransferase homology <STR>

Query Match 38.7%; Score 1265; DB 2; Length 483;  
Best Local Similarity 50.4%; Pred. No. 2.le-86;  
Matches 255; Conservative 79; Mismatches 120; Indels 52; Gaps 9;  
QY 115 LVCITSFISPF--AKDRENARKIHESAGLPFEIFVDAPLNICESRDYKGLYKRARAGEI 172  
DB 3 LLIRSSVYQFHTNRKSPSSSDQTPSKSPFSSFNHNP-----INLYKRNPRWQS 55  
QY 173 KGFTG---IDSDYKPEPPEPVRLTNLSTVSDCVHQVVELLQEQINQVYTIKDIHELFPV 229  
DB 56 LSPFSSMTVKSLIDPDGELV-----ELVV 81  
QY 230 PENKLDHYRAEATLPSITKLDLQWQVVLSEGWATPLKGFMRKEKYLQVNHFDPL-LD 288  
DB 82 PDSETELKKEAESMPKVKLKIDLEWVHVISEGWASPLKGFMRDEYLSLHFNLSRLK 141  
QY 289 DG-VLNMSIPIVLPVSAEDKTRLEGCSEKVF--LAHGRRVAILRDAEYEHKKEKRSRV 345  
DB 142 DGLSNMSELPVLAIDDTKEQI-GVSKNALVSPQGDIVGSLRVEIYKKNKEARIAT 200  
QY 346 WGTTCCKPHI-KVMESGDMLVGGDLQVLEKIRWNGDLDQYRLTPTLELKOKCKEMNADA 404  
DB 201 WGTTSPLPYVEEHITPSSGNLIGGLVLEFQPVKYNVDGLDHYRLSKPLRKEEDNRKADA 260  
QY 405 VFAQLRNPVNHGALLMODTCRRLLERGYKHPVLLHLLPLGQWTKDDDDVPLDWRMKQIAA 464  
DB 261 VFAQLRNPVNHGALLMNDTRKRLLENGYKNPILLLPLGGFTKADDDVPLDVRMEQHSK 320  
QY 465 VLEEGVLDPKSTIVAIFPSPMLYAGTVEQWQHCRSMIAGANFYIVGRDPAGMPHPETKK 524  
DB 321 VLEDGVLDPETIVSIFPSPHYAGTVEQWQHAKARINAGANFYIVGRDPAGMGHPTKGR 380  
QY 525 DLYETHGKGLVSMAPGLTSVEIIPFRVAAYNKAAMDFYDPARHNEFDFTSGTRMKRL 584  
DB 381 DLYDPDHCKGLVSMAPGLEKNILPFRVAAYDTVEKMAFFDPTRAKEFFLFTSGTKMRTY 440  
QY 585 AREGENPPDGPMPAKWAKVLTIDYRS 610  
DB 441 RRTGSPDPDGMCPGSGWNLVKKYES 466

## RESULT 9

T52659  
sulfate adenylyltransferase (EC 2.7.7.4) aps4 precursor [validated] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 03-Nov-2000  
C:Accession: T52659  
R:Hatzfeld, Y.; Lee, S.; Lee, M.; Leustek, T.; Saito, K.  
Gene 248, 51-58, 2000

A:Title: Functional characterization of a gene encoding a fourth ATP sulfurylase isoform  
A:Reference number: Z26165  
A:Accession: T52659  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-469 <HAT>  
A:Cross-references: EMBL:AJ012586; PIDN:CAB42640.1  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: aps4  
A:Genome: nuclear  
C:Function:  
A:Description: EC 2.7.7.4 [validated, MUID:20267847]  
A:Pathway: sulfate assimilation  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: chloroplast; nucleotidyltransferase

Query Match 38.5%; Score 1259.5; DB 2; Length 469;  
Best Local Similarity 60.7%; Pred. No. 5.2e-86;  
Matches 236; Conservative 69; Mismatches 79; Indels 5; Gaps 5;  
QY 227 LFVPEKNKLDHYRAEATLPS-LSITKLDLQWQVVLSEGWATPLKGFMRKEKYLQVNHFDPT 285  
DB 64 LVVEESRRVVKHEAETVPARIKLARVDLEWVHVISEGWASPLKGFMRSEFLQTLHFNS 123  
QY 286 L-LDDG-VLNMSIPIVLPVSAEDKTRLEGCSEKVFIAHG--GRRVAILRDAEYEHKKEERC 342  
DB 124 FRLEDGVSVMNSVPVLAIDDDQKFRIGDSNQVTLVDSVGNPIAILNDIEIYKHPKEERI 183  
QY 343 SRVWGTTCCKPHI-KVMESGDMLVGGDLQVLEKIRWNGDLDQYRLTPTLELKOKCKEMN 401  
DB 184 ARTWGTTLTARGLYAEAITKAGNLIGGLQVLEPIKYNVDGLDFRLSPQLREEFIRRG 243  
QY 402 ADAFAFOLRNPVNHGALLMODTCRRLLERGYKHPVLLHLLPLGQWTKDDDDVPLDWRMKQ 461  
DB 244 ADAFAFOLRNPVNHGALLMODTCRRLLERGYKHPVLLHLLPLGQWTKADDDVPLDWRMKQ 303  
QY 462 HAAVLEEGVLDPKSTIVAIFPSPMLYAGTVEQWQHCRSMIAGANFYIVGRDPAGMPHPPE 521  
DB 304 HEKVLEDGVLDPETTVSIFPSPMLYAGTVEQWQHAKARINAGANFYIVGRDPAGMGHPT 363  
QY 522 TKDLYETHGKGLVSMAPGLTSVEIIPFRVAAYNKAAMDFYDPARHNEFDFTSGTRM 581  
DB 364 EKRDLYDHADHGKGLVSMAPGLEKNILPFRVAAYDKTGKMAFFDPSQDFLFTSGTKM 423  
QY 582 RKLAREGENPPDGPMPAKWAKVLTIDYRS 610  
DB 424 RGLAKKKENPPDGMCPGSGWNLVLDIYDS 452

## RESULT 10

T01204  
sulfate adenylyltransferase (EC 2.7.7.4) - maize  
N:Alternate names: ATP sulfurylase  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 23-Mar-2001  
C:Accession: T01204  
R:Boichi, A.; Petruccio, S.; Foroni, C.; Tenca, G.L.; Ottonello, S.  
submitted to the EMBL Data Library, July 1997  
A:Description: Sulfate permease and ATP sulfurylase mRNA are coordinately modulated  
A:Reference number: Z14260  
A:Accession: T01204  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-489 <BOL>  
A:Cross-references: EMBL:AF016305; NID:g2738749; PIDN:AAB94542.1; PID:g2738750  
A:Experimental source: cultivar Paulo; root  
C:Function:  
A:Description: catalyzes activation of sulfate to adenylylsulfate  
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology  
C:Keywords: nucleotidyltransferase  
F:70-473/Domain: sulfate adenylyltransferase homology <SAT>









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:08:59 ; Search time 24 Seconds  
(without alignments)  
1062.831 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 3271  
Sequence: 1 MSGIKKQKTENQKSTNVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | ID           | Description        |
|------------|--------|-------------|--------|--------------|--------------------|
| 1          | 3236.5 | 98.9        | 614    | 1 PPS2_HUMAN | O95340 h bifunctio |
| 2          | 2985.5 | 91.3        | 621    | 1 PPS2_MOUSE | O88428 m bifunctio |
| 3          | 2599   | 79.5        | 624    | 1 PPS1_MOUSE | O60967 m bifunctio |
| 4          | 2596   | 79.4        | 624    | 1 PPS1_HUMAN | O43252 h bifunctio |
| 5          | 2563   | 78.4        | 624    | 1 PPS1_CAVPO | O54820 c bifunctio |
| 6          | 529.5  | 16.2        | 206    | 1 KAPS_EMENI | Q92203 emericella  |
| 7          | 523    | 16.0        | 202    | 1 KAPS_YEAST | Q02196 saccharomyc |
| 8          | 522.5  | 16.0        | 211    | 1 KAPS_PENCH | Q12657 penicillium |
| 9          | 500    | 15.3        | 199    | 1 CYC2_BACSU | O08735 bacillus su |
| 10         | 497.5  | 15.2        | 293    | 1 KAP2_ARATH | O49196 arabidopsis |
| 11         | 487.5  | 14.9        | 202    | 1 CYC1_BACHD | Q9Kct0 bacillus ha |
| 12         | 476.5  | 14.6        | 200    | 1 CYSC_ECOLI | P23846 escherichia |
| 13         | 476    | 14.6        | 276    | 1 KAP1_ARATH | O43295 arabidopsis |
| 14         | 469.5  | 14.4        | 312    | 1 KAPS_CATRO | O49204 catharanthu |
| 15         | 463    | 14.2        | 215    | 1 CYSC_VIBCH | Q9Kp21 vibrio chol |
| 16         | 456.5  | 14.0        | 614    | 1 CYSN_MYCTU | O10600 m cyn/cysc  |
| 17         | 453    | 13.8        | 633    | 1 NODQ_RHISB | O07309 r nodq bifu |
| 18         | 451.5  | 13.8        | 197    | 1 CYC1_BACSU | O34577 bacillus su |
| 19         | 451    | 13.8        | 632    | 1 NODQ_RHITR | P52978 r nodq bifu |
| 20         | 449    | 13.7        | 646    | 1 NODQ_RHIS3 | P72339 r nodq bifu |
| 21         | 436    | 13.3        | 641    | 1 NODQ_RHIME | P13442 r nodq bifu |
| 22         | 434.5  | 13.3        | 208    | 1 CYSC_BUCAI | P57497 buchnera ap |
| 23         | 429    | 13.1        | 206    | 1 CYC2_BACHD | Q9K7h6 bacillus ha |
| 24         | 425    | 13.0        | 623    | 1 CYSN_XYLFA | O9pd78 x cyn/cysc  |
| 25         | 418    | 12.8        | 546    | 1 SATC_AQUAE | O6d174 a probable  |
| 26         | 415    | 12.7        | 620    | 1 NODQ_AZOB  | P28604 a nodq bifu |
| 27         | 407    | 12.4        | 196    | 1 CYC1_PSEAE | P57702 pseudomonas |
| 28         | 400.5  | 12.2        | 387    | 1 SAT_DEIRA  | P56864 deinococcus |
| 29         | 397    | 12.1        | 490    | 1 MET3_SCHPO | P78937 schizosacch |
| 30         | 396    | 12.1        | 177    | 1 CYSC_SYNY3 | P72940 synechocyst |
| 31         | 384    | 11.7        | 214    | 1 CYC2_PSEAE | P29811 pseudomonas |
| 32         | 374    | 11.4        | 574    | 1 MET3_EMENI | Q12555 emericella  |
| 33         | 371.5  | 11.4        | 389    | 1 SAT2_BACSU | O06736 bacillus su |

#### RESULT 1

| ID | PPS2_HUMAN  | STANDARD; | PRT; | 614 AA. |
|----|---|-----------|------|---------|
| AC | Q95340; Q9UP30; Q9B2L2; Q9P0G6; Q9UHM1; Q9UKD3;   |           |      |         |
| DT | 30-MAY-2000 (Rel. 39, Created)  |           |      |         |
| DT | 15-JUN-2002 (Rel. 41, Last sequence update)   |           |      |         |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update)   |           |      |         |
| DE | Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK-2)                   |           |      |         |
| DE | [includes: sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyltransferase) (SAT) (ATP-sulfurylase): Adenylylsulfate kinase                   |           |      |         |
| DE | (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)  |           |      |         |
| DE | (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-phosphoadenosine-5'-phosphosulfate synthetase)]   |           |      |         |
| GN | PAPS2 OR ATPSK2.  |           |      |         |
| OS | Homo sapiens (Human).   |           |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |           |      |         |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  |           |      |         |
| OX | NCBI_TaxID=9606;  |           |      |         |
| RN | [1]   |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM A).   |           |      |         |
| RC | TISSUE=Fetal cartilage;   |           |      |         |
| RX | MEDLINE=98442651; PubMed=9771708;   |           |      |         |
| RA | Ul Haque M.F., King L.M., Krakow D., Cantor R.M., Rusiniak M.E., Swank R.T., Superti-Furga A., Haque S., Abbas H., Ahmad W., Ahmad M., Cohn D.H.;   |           |      |         |
| RT | "Mutations in orthologous genes in human spondyloepimetaphyseal dysplasia and the brachymorphic mouse."   |           |      |         |
| RL | Nat. Genet. 20:157-162(1998).   |           |      |         |
| RN | [2]   |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM A).   |           |      |         |
| RA | Franson V.L., Gibson M.A., Hatzinikolas G., Cleary E.G., Woolatt E., Sutherland G.R.;   |           |      |         |
| RL | Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.   |           |      |         |
| RN | [3]   |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM A).   |           |      |         |
| RA | Fuda H., Shimizu C., Strott C.A.;   |           |      |         |
| RT | "Human bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase: differential expression of isoforms and effect of polymorphisms on activity." |           |      |         |
| RL | Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.   |           |      |         |
| RN | [4]   |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM A).   |           |      |         |
| RX | MEDLINE=20145452; PubMed=10679223;  |           |      |         |
| RA | Xu Z.-H., Ottensm D.M., Freimuth R.R., Carlini E.J., Wood T.C., Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J., Weinschilboun R.M.;      |           |      |         |
| RT | "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and PAPSS2: gene cloning, characterization and chromosomal localization."        |           |      |         |
| RL | Biochem. Biophys. Res. Commun. 268:437-444(2000).   |           |      |         |
| RN | [5]   |           |      |         |
| RP | SEQUENCE FROM N.A. (ISOFORM B).   |           |      |         |
| RC | TISSUE=Liver;   |           |      |         |
| RX | MEDLINE=20026854; PubMed=10559207;  |           |      |         |
| RA | Kurima K., Singh B., Schwartz N.B.;   |           |      |         |
| RT | "Genomic organization of the mouse and human genes encoding the ATP   |           |      |         |

O66036 chromatium  
P74241 synechocyst  
P56861 deinococcus  
O12650 penicillium  
O34764 bacillus su  
Q9Y872 candida alb  
Q9YCR6 aeropyrum p  
P08536 saccharomyc  
O76156 entamoeba h  
P56863 pyrococcus  
P56862 aspergillus  
Q9YCR4 aeropyrum p

RT sulfurylase/adenosine 5'-phosphosulfate kinase isoform SK2.";  
 RL J. Biol. Chem. 274:33306-33312(1999).  
 RN [6]

RP SEQUENCE FROM N.A. (ISOFORM B).  
 RA Venkatchalam K.V., Fuda H., Strott C.A.;  
 RT "3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]

RP SEQUENCE FROM N.A. (ISOFORM A).  
 RA Strausberg R.;  
 RT Tissue-Colon;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS  
 KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE  
 ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE  
 GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE  
 SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS  
 YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS). ACTIVATED SULFATE DONOR  
 USED BY SULFOTRANSFERASES. IN MAMMALS, PAPS IS THE SOLE SOURCE OF  
 SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-  
 ACTIVATION PATHWAY. MAY HAVE A IMPORTANT ROLE IN SKELETOGENESIS  
 DURING POSTNATAL GROWTH (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.  
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
 phosphoadenylylsulfate.

CC -1- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO  
 THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION  
 OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; are  
 produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN CARTILAGE.

CC -1- DISEASE: DEFECTS IN PAPSS2 ARE THE CAUSE OF SPONDYLOEPHYSEAL  
 DYSPLASIA, PAKISTANI TYPE (SEMD), AN AUTOSOMAL RECESSIVE DISEASE.  
 THIS FORM OF SEMD IS CHARACTERIZED BY A DYSPLASIA THAT IS  
 PRIMARILY EPIPHYSEAL WITH ONLY MILD METAPHYSEAL ABNORMALITIES.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE  
 FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE  
 ADENYLYLSULFATE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AF091242; AAC64583.1; -  
 EMBL; AF074331; AAD38423.1; -  
 EMBL; AF313907; AAK00296.1; -  
 EMBL; AF160509; AAF40307.2; -  
 EMBL; AF160503; AAF40307.2; JOINED.  
 EMBL; AF160504; AAF40307.2; JOINED.  
 EMBL; AF160505; AAF40307.2; JOINED.  
 EMBL; AF160506; AAF40307.2; JOINED.  
 EMBL; AF160507; AAF40307.2; JOINED.  
 EMBL; AF160508; AAF40307.2; JOINED.  
 EMBL; AF173365; AAF12761.1; -  
 EMBL; AF150754; AAF20366.2; -  
 EMBL; BC009894; AAH09894.1; -  
 Genew; HGNC:8604; PAPSS2.  
 MIM; 603005; -  
 InterPro; IPR002891; APS\_kinase.  
 InterPro; IPR002650; ATP\_sulfurylase.  
 Pfam; PF01583; APS\_kinase; 1.  
 Pfam; PF01747; ATP\_sulfurylase; 1.  
 ProDom; PD002350; APS\_kinase; 1.  
 ProDom; PD002381; APS\_sulfurylase; 1.  
 TIGRams; TIGR00455; apsk; 1.  
 Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;  
 KW ATP-binding; Multigene family; Alternative splicing.  
 FT DOMAIN 1 7210 ADENYLYLSULFATE KINASE.

FT DOMAIN 7211 614  
 FT NP\_BIND 49 56  
 FT ACT\_SITE 123 123  
 FT  
 FT SITE 511 515  
 FT VARSPIC 288 288  
 FT CONFLICT 166 166  
 FT CONFLICT 361 361  
 FT CONFLICT 426 426  
 FT CONFLICT 567 567  
 SQ SEQUENCE 614 AA; 69500 MW; 52F4B6D972DDA91E CRC64;

Query Match 98.9%; Score 3236.5; DB 1; Length 614;  
 Best Local Similarity 99.3%; Pred. No. 1.2e-225;  
 Matches 611; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSGIKKOKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGRCGCTVWLTGSGAGKTTISF 60  
 DB 1 MSGIKKOKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGRCGCTVWLTGSGAGKTTISF 60  
 QY 61 ALLEYLVSHAIPCYSLDGDVNRHGLNRNLGFSQDREENIRRIAEVAKLFADAGLVCTIS 120  
 DB 61 ALLEYLVSHAIPCYSLDGDVNRHGLNRNLGFSQDREENIRRIAEVAKLFADAGLVCTIS 120  
 QY 121 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS 180  
 DB 121 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDS 180  
 QY 181 DYKPEPTEPRLVKTNLSTVSDCVHVVQVVELLQEQNIVPTTIKDIHELFPENKLDHVRAE 240  
 DB 181 DYKPEPTEPRLVKTNLSTVSDCVHVVQVVELLQEQNIVPTTIKDIHELFPENKLDHVRAE 240  
 QY 241 AETLPSLSITKLDLQWVQLSEGWATPLKGFMRKEKYLQVMHFDTLLDDGVINNSIPVL 300  
 DB 241 AETLPSLSITKLDLQWVQLSEGWATPLKGFMRKEKYLQVMHFDTLLDDGVINNSIPVL 300  
 QY 301 PVSAEDKTRLEGCSKFVLAHGGRVAILRDAEYFHRKEKERCSEKRVGTTCTKPHIKVMV 360  
 DB 301 PVSAEDKTRLEGCSKFVLAHGGRVAILRDAEYFHRKEKERCSEKRVGTTCTKPHIKVMV 360  
 QY 361 ESCDMLVGGDLQVLEKIRWMDGLDQVRLTPLELKQCKKEMNADAVAFQLRNPVNHGHAL 420  
 DB 361 ESCDMLVGGDLQVLEKIRWMDGLDQVRLTPLELKQCKKEMNADAVAFQLRNPVNHGHAL 420  
 QY 421 LMDQTCRLLERGYKHPVLLHLLPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTTIVAI 480  
 DB 421 LMDQTCRLLERGYKHPVLLHLLPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTTIVAI 480  
 QY 481 FPSPMYAGPTEVQVHCRSRMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLVSMAP 540  
 DB 481 FPSPMYAGPTEVQVHCRSRMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLVSMAP 540  
 QY 541 GLTSVEIIPFRVAAYNKKAKKAMDYDPAHNEFDFISGTRMRKLAREGENPPDGFMAPKA 600  
 DB 541 GLTSVEIIPFRVAAYNKKAKKAMDYDPAHNEFDFISGTRMRKLAREGENPPDGFMAPKA 600  
 QY 601 WKVLTDDYRSEMDKN 615  
 DB 601 WKVLTDDYRS-LEKN 614

## RESULT 2

PSS2\_MOUSE STANDARD; PRT; 621 AA.  
 ID PSS2\_MOUSE  
 AC O88428; Q94274;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS  
 synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK 2)  
 DE [includes: sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl  
 transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase  
 (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)]

| FT   | DOMAIN   | 1  | 7211      | ADENYLYLSULFATE KINASE.                               |
|--|--|--|-----------|---|
| FT   | DOMAIN   | 7212   | 621       | SULFATE ADENYLYLTRANSFERASE.                          |
| FT   | NP_BIND  | 50   | 57        | ATP (POTENTIAL).                                      |
| FT   | ACT_SITE   | 124  | 124       | FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY). |
| FT   | SITE   | 517  | 521       | PP-MOTIF (BY SIMILARITY).                             |
| FT   | VARIANT  | 79   | 79        | G -> R (IN BM; ACTIVITY ABOLISHED).                   |
| FT   | VARIANT  | 109  | 109       | R -> K.   |
| FT   | CONFLICT   | 5  | 5         | S -> F (IN REF. 2).                                   |
| FT   | CONFLICT   | 290  | 294       | MISSING (IN REF. 2).                                  |
| FT   | SEQUENCE   | 621 AA;  | 70259 MD; | DE375F06A79EFAAB CRC64;                               |
| Query Match  |  |  |           |   |
| Best Local Similarity  |  |  |           |   |
| Matches 562; Conservative 18; Mismatches 26; Indels 5; Caps 1; |  |  |           |   |
| QY   | 5  | KKOKTENQOKSTNVVYQAHVHSRNKQVVGTRGGFRGCTVWLTGLSGAGKTTISPALEE     | 64        |   |
| DB   | 6  | KMNHKRDQOKSTNVVYQAHVHSRNKQVVGTRGGFRGCTVWLTGLSGAGKTTISPALEE     | 65        |   |
| QY   | 65   | YLVSHAIPCYSLDGDVNRHGLNRNLGFSQCDREENTRRRTAEVAKLFADAGLVCITSFISP  | 124       |   |
| DB   | 66   | YLVSHAIPCYSLDGDVNRHGLNRNLGFSQCDREENTRRRTAEVAKLFADAGLVCITSFISP  | 125       |   |
| QY   | 125  | FAKDRENARKIHESAGLPEFPEFVDPAPLNCESRDYKGLYKRARAGEIKGFTGIDS DYK   | 184       |   |
| DB   | 126  | FAKDRENARKIHESAGLPEFPEFVDPAPLNCESRDYKGLYKRARAGEIKGFTGIDS DYK   | 185       |   |
| QY   | 185  | PETPERVLKTNLSTVSDCVHVVYVLLQEQNIVPTTIKDIHELFPVENKLDHVRAEATL     | 244       |   |
| DB   | 186  | PETPECVLKTNLSSVSDCVHVVYVLLQEQNIVPTTIKDIHELFPVENKLDHVRAEATL     | 245       |   |
| QY   | 245  | PSLSITPKLDLQWQVLSSEGWATPLKGFMRKEYLQVMHFDTLDD---DGVINMSPIV      | 299       |   |
| DB   | 246  | PSLPITPKLDLQWQVLSSEGWATPLKGFMRKEYLQTLHFDTLDDGVVPRDGVINMSPIV    | 305       |   |
| QY   | 300  | LPVSADCKTRLEGCSKFVLAHGRRVAILRDABEFYHRKEERCERSRVWGTTCTTKHPHIKMV | 359       |   |
| DB   | 306  | LPVSADCKTRLEGCSKFVLAHGRRVAILRDABEFYHRKEERCERSRVWGTTCTTKHPHIKMV | 365       |   |
| QY   | 360  | MESGDWLVGDLQVLEKIRWNCGLDQYRLTPTLELKQCKEMNADAVAFOLRNPVNHGA      | 419       |   |
| DB   | 366  | MESGDWLVGDLQVLEKIRWNCGLDQYRLTPTLELKQCKEMNADAVAFOLRNPVNHGA      | 425       |   |
| QY   | 420  | LLMQDTCRRLLERGYKHPVLLHLPLGCGWTKDDDPVLDWRMKQHAALVEGLVDPKSTIVA   | 479       |   |
| DB   | 426  | LLMQDTCRRLLERGYKHPVLLHLPLGCGWTKDDDPVLDWRMKQHAALVEGLVDPKSTIVA   | 485       |   |
| QY   | 480  | IPSPMLYAGPTEVQWHKCSRMITAGANFYLVGRDPAGMPHPETKKDLYEPETHGKVLMSA   | 539       |   |
| DB   | 486  | IPSPMLYAGPTEVQWHKCSRMITAGANFYLVGRDPAGMPHPETKKDLYEPETHGKVLMSA   | 545       |   |
| QY   | 540  | PGLTSEVILIPFVAAYNKAKMDFYDPARHNEFDFISGTRMRKCLAREGENPPDGFMAPK    | 599       |   |
| DB   | 546  | PGLTSEVILIPFVAAYNKAKMDFYDPARHNEFDFISGTRMRKCLAREGENPPDGFMAPK    | 605       |   |
| QY   | 600  | ANKVLTDYRS   | 610       |   |
| DB   | 606  | ANKVLTDYRS   | 616       |   |
| RESULT 3   |  |  |           |   |
| ID   | PPSL_MOUSE   | STANDARD;  | PPRT;     | 624 AA.   |
| AC   | Q60967;  |  |           |   |
| DT   | 30-MAY-2000  | (Rel. 39, Created)   |           |   |
| DT   | 30-MAY-2000  | (Rel. 39, Last sequence update)                                |           |   |
| DT   | 16-OCT-2001  | (Rel. 40, Last annotation update)                              |           |   |
| DE   | Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SK1) (SK 1)  |  |           |   |
| DE   | [includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylylate transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase) |  |           |   |









entitles requires a license agreement (See <http://www.isb-sib.ch/announcement> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC TIGRFAMS; TIGR00455; apsk; 1.  
CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
CC Phosphorylation.  
CC  
CC EMBL; X57990; CAA41055.1; -  
DR EMBL; S55315; AAB19854.1; -  
DR EMBL; X65124; CAA46252.1; -  
DR EMBL; Z28001; CAA81833.1; -  
DR PIR; S17244; S17244.  
DR SGD; S0001484; MET14.  
DR InterPro; IPR002891; APS\_kinase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR TIGRFAMS; TIGR00455; apsk; 1.  
DR Methionine biosynthesis; Cysteine biosynthesis; Transferase;  
KW Kinase; ATP-binding; Phosphorylation.  
FT NP\_BIND 31 38 ATP (BY SIMILARITY).  
FT ACT\_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
SQ SEQUENCE 202 AA; 23060 MW; 32F3Q1FBG6B2F41D3 CRC64;

Query Match 16.0%; Score 523; DB 1; Length 202;  
Best Local Similarity 56.6%; Pred. No. 8.3e-31;  
Matches 99; Conservative 29; Mismatches 47; Indels 0; Gaps 0;

QY 42 GCTVWLTGLSGAGKTTISFALEEVYVSHAIPCYSLDGDVNRHGLNRNLGFSFGDREINR 101  
DB 24 GCTIWLTLGLSASGKSTIACALEQLLQNLKSAYRLDGNIRFGLNKDLGFSKDRNENIR 83  
QY 102 RTAEVAKLFADAGLVCTITSFISPFADKRENAKRTSAGLPFFEIFVDAPLNICESRDVK 161  
DB 84 RISEVSKLFDASCAISITSFISPYRVDNRARELUKAGLKFTEIFVDVPLEVAEQDPK 143  
QY 162 GLYKRARGEIKGTGIDSDYEKPEPTERVLKTLNLTSLVSDCVHVVLELQEQNIV 216  
DB 144 GLYKKAREGVKEFTGISAPYEAPEAKPELHLRTDQKTEVCATIIYEVLISEXII 198

RESULT 8  
KAPS\_PENCH STANDARD; PRT; 211 AA.  
AC Q12657;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Adenylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-  
DE phosphotransferase)  
DE Penicillium chrysogenum.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eutotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.  
OX NCBI\_TaxID=5076;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 24791;  
RA Foster B.A.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylsulfate = ADP + 3'-  
CC phosphoadenylsulfate.  
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT  
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.  
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U39393; AAA81521.1; -  
DR InterPro; IPR002891; APS\_kinase.

DR TIGRFAMS; TIGR00455; apsk; 1.  
DR Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
DR Phosphorylation.  
DR  
DR EMBL; X57990; CAA41055.1; -  
DR EMBL; S55315; AAB19854.1; -  
DR EMBL; X65124; CAA46252.1; -  
DR EMBL; Z28001; CAA81833.1; -  
DR PIR; S17244; S17244.  
DR SGD; S0001484; MET14.  
DR InterPro; IPR002891; APS\_kinase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR TIGRFAMS; TIGR00455; apsk; 1.  
DR Methionine biosynthesis; Cysteine biosynthesis; Transferase;  
KW Kinase; ATP-binding; Phosphorylation.  
FT NP\_BIND 31 38 ATP (BY SIMILARITY).  
FT ACT\_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
SQ SEQUENCE 206 AA; 23028 MW; 6AADD483E2BCA1CD CRC64;

Query Match 16.2%; Score 529.5; DB 1; Length 206;  
Best Local Similarity 51.4%; Pred. No. 2.9e-31;  
Matches 107; Conservative 33; Mismatches 57; Indels 11; Gaps 3;

QY 15 STNVYQAHVSRNKRQGVQTRGRCITVWLTGLSGAGKTTISFALEEVYVSHAIPCY 74  
DB 2 ATINTHHA-GLTRNERNLRQKQ---KGLTWTGLSAGKSTIAVLEHQLLQRLHAY 56  
QY 75 SLDGDVNRHGLNRNLGFSFGDREINIRIAEVAKLFDAGLVCTITSFISPFADKRENAK 134  
DB 57 RLDDGNVRFGLNKDLGFSADRNENIRIAEVAKLFDADSSIAITSFISFRADRDYARK 116  
QY 135 IHE-----SAGLPFFEIFVDAPLNICESRDVKGLYKRARGEIKGTGIDSDYEKPEPT 188  
DB 117 LHEVTPNDSTGLPVEVFVDVPEVAEAKRDKPKGLYKKAREGIIKEFTGISPYEAPENP 176  
QY 189 ERLVKTNLSTVSDCVHQVVELLQEQNIV 216  
DB 177 EVHVKNVDLPQIEAVKQIIDYLDSSKLL 204

RESULT 7  
KAPS\_YEAST STANDARD; PRT; 202 AA.  
AC Q02196;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Adenylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-  
DE phosphotransferase).  
DE MET14 OR YKL001C.  
GN OS  
OC Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB320;  
RX MEDLINE=91375456; PubMed=1654509;  
RA Korch C., Mountain H.A., Bystrom A.S.;  
RA "Cloning, nucleotide sequence, and regulation of MET14, the gene  
RA encoding the APS kinase of Saccharomyces cerevisiae.";  
RA Mol. Gen. Genet. 229:96-108(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c;  
RX MEDLINE=93070612; PubMed=1441752;  
RA Duesterhoet A., Philippsen P.;  
RA "DNA sequencing and analysis of a 24.7 kb segment encompassing  
RA centromere CEN1 of Saccharomyces cerevisiae reveals nine previously  
RA unknown open reading frames.";  
RL Yeast 8:749-759(1992).  
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylsulfate = ADP + 3'-  
CC phosphoadenylsulfate.  
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT  
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.  
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

```

DR Pfam; PF01583; APS_kinase; 1.
DR PRODOM; PD002350; APS_kinase; 1.
DR TIGRFAMS; TIGR00455; apsk; 1.
KW Transferrase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT ACT_SITE 107 107 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT ACT_SITE 107 107 SIMILARITY).
SQ SEQUENCE 211 AA; 23770 MW; 7DDC4BDA867FE7C2 CRC64;
Query Match 16.0%; Score 522.5; DB 1; Length 211;
Best Local Similarity 50.0%; Pred. No. 9.5e-31;
Matches 105; Conservative 35; Mismatches 59; Indels 11; Gaps 3;
QY 15 STNVVQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFALBEYLVS-HPAIPC 73
DB 2 STNITHASALTSERTELNRQ---RGLTWLTGLSAGSKSPLAVELEHQLVDRRVHA 57
QY 74 YSLDGNVHRLNRLNGLFSGDRENNIRIAEVAKLFDAGLVCTISFISPPFAKDRENA 133
DB 58 YRLDGNIRFLNKLGLFSEADRENIRIAEVAKLFDADNSNIAISFISPYRKDRDAR 117
QY 134 KIHESA-----GLPFEFIVDAPLNICESRDVKGLYKKRAGEIKGFTGDSYKPE 187
DB 118 QLHEVATPGEETGLPFEVVDVPEVAGQDPKGLYKKRAGEIKGFTGISAPYEAPAN 177
QY 188 PERVLKTNLSTVSDCVHQVVELLQEQNIYP 217
DB 178 PEVHVKNYELPVQDAVKQIIDYLDTKGYLP 207
RESULT 9
CYC2_BACSU STANDARD; PRT; 199 AA.
AC O06735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable adenylyltransferase kinase (EC 2.7.1.25) (APS kinase) (adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase).
GN YISZ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=98015416; PubMed=9353932;
RA Roche B., Autret S., Levine A., Vannier F., Medina N., Seror S.J.;
RT "A Bacillus subtilis chromosome segment at the 100 degrees to 102
RT degrees position encoding 11 membrane proteins.";
RL Microbiology 143:3309-3312(1997).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borliss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Hollsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Toris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinost S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

```

```

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche M., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Taseato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RA "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y09476; CAA70655.1; ALT_INIT.
DR EMBL; Y99109; CAA12931.1; -.
DR Subtilist; BG13105; yisz.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR PRODOM; PD002350; APS_kinase; 1.
DR TIGRFAMS; TIGR00455; apsk; 1.
KW Transferrase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 34 41 ATP (BY SIMILARITY).
FT ACT_SITE 108 108 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT ACT_SITE 108 108 SIMILARITY).
SQ SEQUENCE 199 AA; 22304 MW; 90EC37FE4B02A123 CRC64;
Query Match 15.3%; Score 500; DB 1; Length 199;
Best Local Similarity 47.0%; Pred. No. 3.7e-29;
Matches 94; Conservative 43; Mismatches 57; Indels 6; Gaps 2;
QY 17 NVVYQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFALBEYLVS-HPAIPC 76
DB 6 NIIHHPAALSKSDQSL---NGHKSCVLWFTGLSGSKSVLANAYDEKLYRKGIQSYVL 61
QY 77 DGDNVHRLNRLNGLFSGDRENNIRIAEVAKLFDAGLVCTISFISPPFAKDRENAKTH 136
DB 62 DGDNTRHGLNKLGLFTQDRIENIRIGEVAKLFDVSGOMILTAFTISPFREDMDVRAFL 121
QY 137 ESAGLPPFFEIFVDAPLNICESRDVKGLYKKRAGEIKGFTGDSYKPEKTPERVILKTNL 196
DB 122 PKG--EFFEIVKCPHVCQDPKGLYKKRAGEIKGFTGDSYKPEALPSDPFIIESDQ 179
QY 197 STVSDCVHQVVELLQEQNIYP 216
DB 180 TSISDGDADLIINALQNRGII 199
RESULT 10
KAP2_ARATH STANDARD; PRT; 293 AA.
AC O49196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS

```

DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-  
 GN phosphosulfate 3'-phosphotransferase).  
 OS AKN2 OR AT4G39940 OR T5J17.110.  
 OC Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Schiffmann S., Schwenn J.-D.;  
 RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase  
 RT (EC2.7.1.25) from *Catharanthus roseus* and an isoform (akn2) from  
 RT Arabidopsis.";  
 RL (1n) Plant Gene Register PCR98-116.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=20083489; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,  
 RA Harris B., Ansong W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Mach R., Mueller M.,  
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohelsel J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
 RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., de Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,  
 RA Pettett A., Rajadream M.A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Fartmann B., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,  
 RA Neumann S., Argirou A., Vitale D., Liquori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schmal S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Cheford F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,  
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott K., Scott K., Johnson D.,  
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Splich J., Ryan E., Andrews S., Geisel C., Layman D.,  
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,  
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,  
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:769-777(1999).  
 CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
 CC phosphoadenylylsulfate.  
 CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS  
 CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE  
 CC BIOSYNTHETIC PATHWAY.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast. (By similarity).

CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF043351; AAC39520.1; -  
 DR EMBL: AL035708; CAB38907.1; -  
 DR EMBL: AL161596; CAB80657.1; -  
 DR InterPro: IPR002891; APS.Kinase.  
 DR Pfam: PF01583; APS.Kinase; 1.  
 DR ProDom: PD002350; APS.Kinase; 1.  
 DR TIGRFAMs: TIGR00455; apsk; 1.  
 KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
 KW Phosphorylation; Transient peptide; Chloroplast.  
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
 FT CHAIN 293 ADENYLYLSULFATE KINASE 2.  
 FT NP\_BIND 142 149 ATP (POTENTIAL).  
 FT ACT\_SITE 196 196 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 293 AA; 31977 MW; 71F1AD6E9B026886 CRC64;  
 Query Match 15.2%; Score 497.5; DB 1; Length 293;  
 Best Local Similarity 48.2%; Pred. No. 9.5e-29;  
 Matches 105; Conservative 33; Mismatches 63; Indels 17; Gaps 4;  
 QY 4 IKKQTNQKSTNNVYQAHVSRKRGQVVGTRGGFRGCTVWLTGSLGAGKTTISPALE 63  
 DB 81 LKEGRNGKKAENIVWHSSICRQQLLQ-----KGVVWITGSGSGKSTVACALS 136  
 QY 64 EYLVSHAIPCVSLDGNVRHGLNRLGSPGDRENTIRRAEVAKLFADAGLVCTTSPIS 123  
 DB 137 KALEFERGLYTLDDGNVRHGLNRLDTKAEHRTENRRIGEVAKLFADVGVCIASLIS 196  
 QY 124 PFAKDRENARKIHESAGLP---PFEIFVDAPLNICESRDVKGLYKRAGAEIKGFTGIDS 180  
 DB 197 PYRRDRDACRSL-----LPDGDVEFVEMDVPVHCVESRDPKGLYKLARAGIKGFTGIDD 251  
 QY 181 DYKPEPTPVRVLTNLTSTVSDCVHVVVLEQNVIPY 218  
 DB 252 PYEAPVCEVYLVKHTGDDESCSPROMAE-----NIISY 284  
 RESULT 11  
 ID CYCL\_BACHD STANDARD; PRT; 202 AA.  
 AC Q9KCT0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
 DE 5-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-  
 GN BHL489.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OC NCBI\_TaxID=86665;  
 RX [1]  
 RX SEQUENCE FROM N.A.  
 RX STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

phosphoadenylylsulfate.  
-!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE BIOSYNTHETIC PATHWAY.  
-!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AP001512; BAB05208.1; -  
DR InterPro: IPR002891; APS\_kinase.  
DR Pfam: PF01583; APS\_kinase; 1.  
DR ProDom: PD002350; APS\_kinase; 1.  
DR TIGRFAMs: TIGR00455; apsk; 1.  
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
KW Phosphorylation; Complete proteome.  
FT NP\_BIND 36 43 ATP (BY SIMILARITY).  
FT ACT\_SITE 110 110 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).  
SQ SEQUENCE 202 AA; 22658 MW; 6BB294F9242F1FCD CRC64;  
  
Query Match 14.9%; Score 487.5; DB 1; Length 202;  
Best Local Similarity 48.8%; Pred. No. 3e-28;  
Matches 100; Conservative 31; Mismatches 65; Indels 9; Gaps 3;  
  
QY 9 TENQKSTNVVYQAHVSRNKRQVVTGRTGGFCGVWLTGLSGACKTITISFALEELYVS 68  
DB 3 TSNQ--PHIVWHEASVSKEEROK----RNRUKSCVWFTGLSGSGKSTLANALDKLFE 55  
  
QY 69 HAIPCYSLQDGNVRHGLNRLGSPGDREINIRIAEVAKLFDAGLVCITTSIPSPFXD 128  
DB 56 EGLHSVLDGDNIRHGLNAGLGFSEEDRKNIRIGEVAKLFDVAGVVTAFISFRED 115  
  
QY 129 RENARKIHESAGLPFEIFVDPLNICESRDVGLYKRRAGEIKGFTGIDSDYKEPFP 188  
DB 116 RDNVRGILDG--EFTEVVRCPLETCEKRDPRGLYKARSQDIPFTGTSPPYEPVNP 173  
  
QY 189 ERVLKTNLSTVSCVHQVVELLOEQ 213  
DB 174 ELIIDTQDLAVEAVEKIYALHAQ 198  
  
RESULT 12  
ID CYSC\_ECOLI STANDARD; PRT; 200 AA.  
AC F23846; Q59389; Q59376; P78105; -  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).  
DE CYSC OR B2750 OR Z4058 OR ECS3604.  
GN Escherichia coli, and  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562, 83334;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.  
RX STRAIN=K12;  
RC MEDLINE=92268080; PubMed=1316900;  
RA Leyh T.S., Vogt T.F., Suo Y.;  
RT "The DNA sequence of the sulfate activation locus from Escherichia coli K-12";  
RL J. Biol. Chem. 267:10405-10410(1992).  
RN [2]  
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ACTIVE SITE.

RX MEDLINE=93075778; PubMed=1332767;  
RT Sathishchandran C., Hickman Y.N., Markham G.D.;  
RA "Characterization of the phosphorylated enzyme intermediate formed in the adenosine 5'-phosphosulfate kinase reaction";  
RT Biochemistry 31:11684-11688(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12";  
RL Science 277:1453-1474(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kuhara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12";  
RL DNA Res. 8:11-22(2001).  
RN [6]  
RP CHARACTERIZATION.  
RC MEDLINE=88115389; PubMed=2828368;  
RA Leyh T.S., Taylor J.C., Markham G.D.;  
RT "The sulfate activation locus of Escherichia coli K12: cloning, genetic, and enzymatic characterization";  
RL J. Biol. Chem. 263:2409-2416(1988).  
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-phosphoadenylylsulfate.  
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE BIOSYNTHETIC PATHWAY.  
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M74586; AAA23647.1; -  
DR EMBL; M86936; AAA23503.1; -  
DR EMBL; U29579; AAA69260.1; -  
DR EMBL; AE000358; AAC7592.1; -  
DR EMBL; AE005502; AAG57857.1; -  
DR EMBL; AP002562; BAB37027.1; -  
DR PIR; JN0328; JN0328.  
DR PIR; A44200; A44200.  
DR EcoGene; EG10185; cysC.  
DR InterPro; IPR002891; APS\_kinase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.



DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-  
DE phosphosulfate 3'-phosphotransferase).  
GN AKN.  
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;  
OC Viscaceae; Catharanthus.  
OX NCBI\_TaxID=4058;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP Schiffmann S., Schwenn J.-D.:  
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase  
RT (EC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from  
RT Arabidopsis.";  
RL (In) Plant Gene Register PGR98-116.  
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS  
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE  
CC BIOSYNTHETIC PATHWAY.  
CC -1- SUBCELLULAR LOCATION: Chloroplast (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF044285; AAC31145.1; -  
CC InterPro: IPR002891; APS\_kinase.  
CC Pfam: PF01583; APS\_kinase; 1.  
CC ProDom: PD002350; APS\_kinase; 1.  
CC TIGRFAMS: TIGR00455; apsk; 1.  
CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
KW Phosphorylation; transist peptide; Chloroplast.  
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
FT CHAIN ? 312 ADENYLYLSULFATE KINASE.  
FT NP\_BIND 142 149 ATP (POTENTIAL).  
FT ACT\_SITE 216 216 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
SQ SEQUENCE 312 AA; 33656 MW; 16BBD1FB4B1FE27 CRC64;  
Query Match 14.4%; Score 469.5; DB 1; Length 312;  
Best Local Similarity 47.7%; Pred. No. 1.1e-26;  
Matches 103; Conservative 30; Mismatches 68; Indels 15; Gaps 4;  
QY 3 GIKKQKTENQOKSTNVYQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFAL 62  
Db 100 GKILQTTTGVNSTNLHMKCAVEKSEROEPLQO----RCGVITWITGLSGSKSTLACAL 155  
QY 63 EYLVSHAI PCYSLDGDVNRHGLNRLGSPGDEENIRIAEVAKLFADAGLVCITSF 122  
Db 156 SRGLHAGKLT YLDDGDNVRHGLNRLGSPGDEENIRIAEVAKLFADAGVICIASLI 215  
QY 123 SPFAKREARKITHESAGLP---FFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGID 179  
Db 216 SPYKPPDACSRL-----LPEGDFIEVMDVPLKVCARDPKGLYKLARAGKIKGFTGID 270  
QY 180 SDYEKPEPTEPVLUKTNL-----STVSDCVHQVVELLQE 212  
Db 271 DPVEPPKSEIVLHQKMGCDSPCDLADIVISYLEE 306  
RESULT 15  
CYSC\_VIBCH STANDARD; PRT; 215 AA.  
AC 09KP21;  
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-  
DE phosphotransferase).  
GN CYSC OR VC2558.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=El Tor N16961 / Serotype O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RL Nature 406:477-483(2000).  
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
CC phosphoadenylylsulfate.  
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS  
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE  
CC BIOSYNTHETIC PATHWAY.  
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AE004324; AAF95699.1; -  
CC TIGR: VC2558; -  
CC InterPro: IPR002891; APS\_kinase.  
CC Pfam: PF01583; APS\_kinase; 1.  
CC ProDom: PD002350; APS\_kinase; 1.  
CC TIGRFAMS: TIGR00455; apsk; 1.  
CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;  
KW Phosphorylation; Complete proteome.  
FT NP\_BIND 46 53 ATP (BY SIMILARITY).  
FT ACT\_SITE 120 120 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY  
FT SIMILARITY).  
SQ SEQUENCE 215 AA; 23906 MW; 83B9EE2F295CDDF6 CRC64;  
Query Match 14.2%; Score 463; DB 1; Length 215;  
Best Local Similarity 47.6%; Pred. No. 1.9e-26;  
Matches 100; Conservative 33; Mismatches 65; Indels 12; Gaps 3;  
QY 10 ENQOKSTNVYQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFAL 69  
Db 11 EODAKPENVMHHRHVDKQAKR-----ATLKQORPAVLWFTGLSGAGKSTVAGALENRLAAL 66  
QY 70 AIPCYSLDGDVNRHGLNRLGSPGDEENIRIAEVAKLFADAGLVCITSFISPAKOR 129  
Db 67 GYHTYLLDDGDNVRHGLNRLGSPGDEENIRIRIGELAKLMDAGLIVLTAFTISPHRAER 136  
QY 130 ENARKITHESAGLP---FFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDSVEKPE 186  
Db 127 QMVRDL-----LPNGFELEVYVNTSLDVCARDPKGLYKKARAGEIQFTGIDSAYEAPL 181  
QY 187 IPERVLKTNLSTVSDCVHQVVELLQEQNV 216  
Db 182 NFDIDLPAKEKSVDELVAQCLQALAEHRHII 211

Thu Jun 5 08:31:30 2003

us-09-898-165b-7.rsp

Page 13

Search completed: May 30, 2003, 12:18:02  
Job time : 25 secs

**THIS PAGE BLANK (US70)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:14:14 ; Search time 88 Seconds  
(without alignments)  
1439.988 Million cell updates/sec

Title: US-09-898-165B-7  
Perfect score: 3271  
Sequence: 1 MSGIKKQKTEKQKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhnc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertibrate.\*
- 14: sp.unclassified.\*
- 15: sp.rvirus.\*
- 16: sp.bacteriap.\*
- 17: sp.archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Match | Length | DB | ID     | Description         |
|------------|--------|-------|--------|----|--------|---------------------|
| 1          | 3020.5 | 92.3  | 620    | 11 | Q9JK86 | Q9JK86 cavia porce  |
| 2          | 2965.5 | 90.7  | 621    | 11 | Q9QXS0 | Q9QXS0 mus musculus |
| 3          | 2638.5 | 80.7  | 613    | 13 | Q9QXY2 | Q9QXY2 fugu rubrip  |
| 4          | 2602   | 79.5  | 624    | 4  | Q96TF4 | Q96TF4 homo sapien  |
| 5          | 2596   | 79.4  | 624    | 4  | Q9P1P9 | Q9P1P9 homo sapien  |
| 6          | 2588   | 79.1  | 624    | 4  | Q9UE98 | Q9UE98 homo sapien  |
| 7          | 2367   | 72.4  | 610    | 5  | Q27128 | Q27128 urechis cau  |
| 8          | 2233.5 | 68.3  | 618    | 5  | Q9NDP8 | Q9NDP8 ciona intes  |
| 9          | 2039   | 62.3  | 629    | 5  | Q961A8 | Q961A8 drosophila   |
| 10         | 2031.5 | 62.1  | 630    | 5  | Q9VW48 | Q9VW48 drosophila   |
| 11         | 1887   | 57.7  | 652    | 5  | Q22501 | Q22501 caenorhabdi  |
| 12         | 1351.5 | 41.3  | 265    | 4  | Q9UIR2 | Q9UIR2 homo sapien  |
| 13         | 1301   | 39.8  | 463    | 10 | Q9LIK9 | Q9LIK9 arabidopsis  |
| 14         | 1299   | 39.7  | 424    | 10 | Q43170 | Q43170 solanum tub  |
| 15         | 1282   | 39.2  | 465    | 10 | Q42520 | Q42520 arabidopsis  |
| 16         | 1282   | 39.2  | 465    | 10 | Q96530 | Q96530 arabidopsis  |

|    |        |      |     |    |        |                     |
|----|--------|------|-----|----|--------|---------------------|
| 17 | 1282   | 39.2 | 465 | 10 | O23324 | O23324 arabidopsis  |
| 18 | 1274   | 38.9 | 463 | 10 | Q43183 | Q43183 solanum tub  |
| 19 | 1268.5 | 38.8 | 476 | 10 | Q43870 | Q43870 arabidopsis  |
| 20 | 1265   | 38.7 | 483 | 10 | Q96541 | Q96541 brassica ol  |
| 21 | 1264   | 38.6 | 461 | 10 | Q92N28 | Q92N28 brassica ju  |
| 22 | 1260   | 38.5 | 483 | 10 | Q9SBL0 | Q9SBL0 brassica ol  |
| 23 | 1259.5 | 38.5 | 469 | 10 | Q9S7D8 | Q9S7D8 arabidopsis  |
| 24 | 1257.5 | 38.4 | 470 | 10 | Q92N29 | Q92N29 brassica ju  |
| 25 | 1254.5 | 38.4 | 476 | 10 | Q8RWJ3 | Q8RWJ3 arabidopsis  |
| 26 | 1252.5 | 38.3 | 458 | 10 | Q8W1X3 | Q8W1X3 allium cepa  |
| 27 | 1250   | 38.2 | 463 | 10 | Q9SE02 | Q9SE02 arabidopsis  |
| 28 | 1245   | 38.1 | 463 | 10 | Q42519 | Q42519 arabidopsis  |
| 29 | 1244.5 | 38.0 | 461 | 10 | Q9SDP4 | Q9SDP4 allium cepa  |
| 30 | 1230.5 | 37.6 | 489 | 10 | Q48888 | Q48888 zea mays (m  |
| 31 | 1230   | 37.6 | 476 | 10 | Q92WM0 | Q92WM0 oryza sativ  |
| 32 | 1212.5 | 37.1 | 459 | 10 | Q96349 | Q96349 brassica na  |
| 33 | 1158.5 | 35.4 | 336 | 5  | Q95P41 | Q95P41 aedes aegypt |
| 34 | 1108.5 | 33.9 | 461 | 10 | Q8SAG1 | Q8SAG1 glycine max  |
| 35 | 556    | 17.0 | 202 | 3  | Q9P7G9 | Q9P7G9 schizosacch  |
| 36 | 551.5  | 16.9 | 200 | 16 | Q97MT8 | Q97MT8 clostridium  |
| 37 | 524    | 16.0 | 202 | 3  | Q9HGF8 | Q9HGF8 saccharomyc  |
| 38 | 524    | 16.0 | 202 | 3  | Q9C2Y6 | Q9C2Y6 saccharomyc  |
| 39 | 508.5  | 15.5 | 201 | 2  | Q9EY11 | Q9EY11 klebsiella   |
| 40 | 493.5  | 15.1 | 213 | 16 | Q8ZBF3 | Q8ZBF3 yersinia pe  |
| 41 | 481.5  | 14.7 | 201 | 16 | Q8XF34 | Q8XF34 salmonella   |
| 42 | 481.5  | 14.7 | 635 | 16 | Q9A882 | Q9A882 caulobacter  |
| 43 | 472.5  | 14.4 | 208 | 10 | Q9SRW7 | Q9SRW7 arabidopsis  |
| 44 | 460.5  | 14.1 | 290 | 10 | Q9FJX1 | Q9FJX1 arabidopsis  |
| 45 | 452    | 13.8 | 645 | 16 | Q98504 | Q98504 rhizobium 1  |

#### ALIGNMENTS

#### RESULT 1

| ID | Q9JK86  | PRELIMINARY; | PRT; | 620 AA. |
|----|---|--------------|------|---------|
| AC | Q9JK86;   |              |      |         |
| DT | 01-OCT-2000 (Tremblrel_15, Created)   |              |      |         |
| DT | 01-OCT-2000 (Tremblrel_15, Last sequence update)  |              |      |         |
| DT | 01-JUN-2002 (Tremblrel_21, Last annotation update)  |              |      |         |
| DE | Adenosine 5'-phosphosulfate kinase/ATP sulfurylase 2.   |              |      |         |
| OS | Cavia porcellus (Guinea pig).   |              |      |         |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                                 |              |      |         |
| OC | Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.                                    |              |      |         |
| OX | NCBI_TaxID=10141;   |              |      |         |
| RN | [1]   |              |      |         |
| RP | SEQUENCE FROM N.A.  |              |      |         |
| RA | Fuda H., Strott C.A.;   |              |      |         |
| RT | "Guinea pig bifunctional adenosine 5'-phosphosulfate kinase/ATP sulfurylase 2, gpAPS synthase 2." |              |      |         |
| RL | Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.   |              |      |         |
| DR | EMBL; AF251798; AAF70194.1;   |              |      |         |
| DR | InterPro; IPR002891; APS_kinase.  |              |      |         |
| DR | InterPro; IPR002650; ATP_sulfurylase.   |              |      |         |
| DR | Pfam; PF01583; APS_kinase; 1.   |              |      |         |
| DR | Pfam; PF01747; ATP_sulfurylase; 1.  |              |      |         |
| DR | ProDom; PD002350; APS_kinase; 1.  |              |      |         |
| DR | ProDom; PD002381; APS_kinase; 1.  |              |      |         |
| DR | TIGRFAMs; TIGR00455; apsk; 1.   |              |      |         |
| KW | Kinase.   |              |      |         |
| SQ | SEQUENCE 620 AA; 70031 MW; 50E187D16335985A CRC64;  |              |      |         |

|                       |                  |  |           |             |
|-----------------------|------------------|--|-----------|-------------|
| Query Match           | 92.3%;           | Score 3020.5;  | DB 11;    | Length 620; |
| Best Local Similarity | 91.5%;           | Pred. No. 9.2e-235;                                      |           |             |
| Matches 563;          | Conservative 26; | Mismatches 21;   | Indels 5; | Gaps 1;     |
| QY                    | 1                | MSGIKKQKTEKQKSTNVVYQAHVSRNRKRGQVGTGRGCTCTVMTGLSGAGKTTISF | 60        |             |
| DB                    | 1                | MSGVKKQKTESQKSTNVVYQAHVSRNRKRGQVGTGRGCTCTVMTGLSGAGKTTISF | 60        |             |
| QY                    | 61               | ALEEVLVSHAPCYSLDGDNVHGLNRNLGFGPDREENIRIAEVAKIFADAGLCVCTS | 120       |             |

2

2

DE 3'-phosphoadenosine 5'-phosphosulfate synthase 2.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21455682; PubMed=11571655;  
RA Yu W.P., Pallen C.J., Tay A., Jirik F.R., Brenner S., Tan Y.H.,  
RA Venkatesh B.;  
RT "Conserved synteny between the Fugu and human PTEN locus and the  
RT evolutionary conservation of vertebrate PTEN function.";  
RL Oncogene 20:5554-5561(2001).  
DR EMBL; AF325922; AAL08416.1; -  
DR InterPro; IPR002891; APS\_kinase.  
DR InterPro; IPR002650; ATP-sulfurylase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfurylase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR ProDom; PD002381; ATP-sulfurylase; 1.  
DR TIGRFAMs; TIGR00455; apsk; 1.  
SQ SEQUENCE 613 AA; 68948 MW; 28E50148377C4169 CRC64;  
  
Query Match 80.7%; Score 2638.5; DB 13; Length 613;  
Best Local Similarity 79.3%; Pred. No. 5.6e-204;  
Matches 487; Conservative 60; Mismatches 66; Indels 1; Gaps 1;  
  
QY 1 MSGIKKOTENOOKSTNNVYQAHVSRNKRKGVVTRGFRGCTTWTGLSGAGTTISF 60  
DB 1 MSGAKKLT-NLERSTNNVYQAHVSRNKRKGVVTRGFRGCTTWTGLSGAGTTISF 59  
  
QY 61 ALEEVYVSHAIPCYSLDGNVRHGLNRLGSPGDRNIRIAEVAKLADAGLVCITS 120  
DB 60 ALEEVYVSHAIPCYSLDGNVRHGLNRLGSPGDRNIRIAEVAKLADAGLVCITS 119  
  
QY 121 FISPAKRENAKTHESAGLPFFELFVDAPLNICESRDVKGLYKRAGEIKGTGTDS 180  
DB 120 FISPAKRENAKTHESAGLPFFELFVDAPLNICESRDVKGLYKRAGEIKGTGTDS 179  
  
QY 181 DYKPEKTEPVLKTLNLTVDVSCVHVVQLQONIVPVYTIKDIHELFPENKLDHVAE 240  
DB 180 NYESPDRLVLKLTGELVDECLQVLELRENDILPTGIMEEINELFVAENKLSAAVAE 239  
  
QY 241 AETPLSLITKLDLQVQVLSGWTAPLKGFMRKEYLQVMHFDTLDDGVINMSIPIVL 300  
DB 240 ASTLPTISTTKLDLQVQVLSGWTAPLKGFMRKEYLQVMHFDTLDDGVINMSIPIVL 299  
  
QY 301 PVSADKTRLECCSFVLAHGRVAILRDABFEYHRKEERCSRVTGCTTKPHIKVMW 360  
DB 300 PVSNETKOLEGCEAVTLEYQGSRAILLRNMEFYANRIEERCARQWGTCPQHPYIKWM 359  
  
QY 361 ESGDWLVGDLQVLEKIRWNDGLDOYRLTPTLELKOKCKEMNADAFATOLRNPVINGHAL 420  
DB 360 EGGDWLVGDLQVLEKIRWNDGLDOYRLTPTLELKOKCKEMNADAFATOLRNPVINGHAL 419  
  
QY 421 LMQDTCRLRLRGYKHPVLLHPLGSGWTKDDVPLDWRMKQAAVLEECVLPDKSTIVAI 480  
DB 420 LMQDTCRLRLRGYKHPVLLHPLGSGWTKDDVPLDWRMKQAAVLEECVLPDKSTIVAI 479  
  
QY 481 FSPMLYAGPTEVQVHCRSRMTAGANFYVGRDPAGMPHPETPKKOLYEPHGGKVLNAP 540  
DB 480 FSPMLYAGPTEVQVHCRSRMTAGANFYVGRDPAGMPHPETPKKOLYEPHGGKVLNAP 539  
  
QY 541 GLTSVEIIPFVAANKAKMDYDPAHNEFDELSGTRMKLAREGENPPDGFMAPKA 600  
DB 540 GLTSVEIIPFVAANKAKMDYDPAHNEFDELSGTRMKLAREGENPPDGFMAPKA 599  
  
QY 601 KVLVTDYRSEMDK 614  
DB 600 KVLVTDYRSEMDK 613

RESULT 4  
Q96TF4 PRELIMINARY; PRT; 624 AA.  
AC Q96TF4; Q96FBI;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.  
GN PAPSS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20145452; PubMed=10679223;  
RA Xu Z.-H., Otterness D.M., Frelmuth R.R., Carlini E.J., Wood T.C.,  
RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,  
RA Weinsilboum R.M.;  
RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSSI) and  
RT PAPSS2: gene cloning, characterization and chromosomal localization.";  
RL Biochem. Biophys. Res. Commun. 268:437-444(2000).  
RN [2]  
RP SEQUENCE OF 22-624 FROM N.A.  
RC TISSUE=EYE;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF097721; AAF40235.1; -  
DR EMBL; AF097710; AAF40235.1; JOINED.  
DR EMBL; AF097711; AAF40235.1; JOINED.  
DR EMBL; AF097712; AAF40235.1; JOINED.  
DR EMBL; AF097713; AAF40235.1; JOINED.  
DR EMBL; AF097714; AAF40235.1; JOINED.  
DR EMBL; AF097715; AAF40235.1; JOINED.  
DR EMBL; AF097716; AAF40235.1; JOINED.  
DR EMBL; AF097717; AAF40235.1; JOINED.  
DR EMBL; AF097718; AAF40235.1; JOINED.  
DR EMBL; AF097719; AAF40235.1; JOINED.  
DR EMBL; AF097720; AAF40235.1; JOINED.  
DR EMBL; BC011392; AAH11392.1; -  
DR InterPro; IPR002891; APS\_kinase.  
DR InterPro; IPR002650; ATP-sulfurylase.  
DR Pfam; PF01583; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfurylase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR ProDom; PD002381; ATP-sulfurylase; 1.  
DR TIGRFAMs; TIGR00455; apsk; 1.  
SQ SEQUENCE 624 AA; 70833 MW; A3DC9B9943568CDD6 CRC64;  
  
Query Match 79.5%; Score 2602; DB 4; Length 624;  
Best Local Similarity 77.1%; Pred. No. 5.1e-201;  
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;  
  
QY 5 KQKQTEK-----QKSTNNVYQAHVSRNKRKGVVTRGFRGCTTWTGLSGAGTTI 58  
DB 9 KVKLSNNAQNMQRATNVTYQAHVSRNKRKGVVTRGFRGCTTWTGLSGAGTTV 68  
  
QY 59 SFALIEYVSHAIPCYSLDGNVRHGLNRLGSPGDRNIRIAEVAKLADAGLVC 118  
DB 59 SFALIEYVSHAIPCYSLDGNVRHGLNRLGSPGDRNIRIAEVAKLADAGLVC 128  
  
QY 119 TSFISPAKRENAKTHESAGLPFFELFVDAPLNICESRDVKGLYKRAGEIKGTGT 178  
DB 119 TSFISPAKRENAKTHESAGLPFFELFVDAPLNICESRDVKGLYKRAGEIKGTGT 188  
  
QY 179 DSDYKPEKTEPVLKTLNLTVDVSCVHVVQLQONIVPVYTIKDIHELFPENKLDHVR 238  
DB 179 DSDYKPEKTEPVLKTLNLTVDVSCVHVVQLQONIVPVYTIKDIHELFPENKLDHVR 248  
  
QY 239 AEAETPLSLITKLDLQVQVLSGWTAPLKGFMRKEYLQVMHFDTLDDGVINMSIPI 298  
DB 239 AEAETPLSLITKLDLQVQVLSGWTAPLKGFMRKEYLQVMHFDTLDDGVINMSIPI 308

|  |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
|--|---|---|---|-----------------------|---------|--|--|--|--|--|--|--|--|--|
| QY   | 299   | VLPYSAEDKTRLEGC                         | KFVLAHGRRVAILRDAEYEHKREERC              | SRVGGTTCTKHPIHKM      | 358     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 309   | VLTATHEDKE                              | RLDCTAFALMYEGSRVAILRNPEFEH              | KREERCARQWGTCKNHPYIKM | 368     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| QY   | 359   | VMESGDWLVGGD                            | LVQVLEKIRWNGDLDQVRLPLELKQKCKEMNADAVAFQ  | LRLNVPVHGH            | 418     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 369   | VMEQGDWLVGGD                            | LVQVLDVRYVNDGLDQVRLPELTKQKFKDMNADAVAFQ  | LRLNVPVHGH            | 428     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| QY   | 419   | ALLMQDTCRRLLE                           | RGYHHPVLLLLHPLGGWTKDDVPLDWRMKQHA        | AVLEEGVLDPKSTIV       | 478     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 429   | ALLMQDTHKQLLE                           | RGYRRPVLLLLHPLGGWTKDDVPLMWRMKQHA        | AVLEEGVLDPETTVV       | 488     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| QY   | 479   | AIFSPMLYAGPTE                           | VQWCHRCRMTAGANFYIVGRDPAGMHPETK          | KDLYEPTHGKVL          | 538     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 489   | AIFSPMMYAGPTE                           | VQWCHRCARMVAGANFYIVGRDPAGMHPETG         | KDLYEPTHGAKVL         | 548     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| QY   | 539   | APGTSVEIIPRVA                           | AYNAKKAKKADFYDPPARHNEFDFISGTRMRKLAREGEN | PPDGGFMAP             | 598     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 549   | APGLITLIEIPRVA                          | AYNAKKKKRMDYDSEHHEDFEISGTRMRKLAREGQ     | KPPGEGFMAP            | 608     |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| QY   | 599   | KAKVLTDDYYS                             | 610                                     |                       |         |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| Db   | 609   | KAWTVLTYEYKS                            | 620                                     |                       |         |  |  |  |  |  |  |  |  |  |
|  |   |   | :         :                             | :             :       |         |  |  |  |  |  |  |  |  |  |
| RESULT 5   |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| Q9P1P9   |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| ID   | Q9P1P9  | PRELIMINARY:                            |   | PRT:                  | 624 AA. |  |  |  |  |  |  |  |  |  |
| AC   | Q9P1P9  |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| DT   | 01-OCT-2000   | (TrEMBLrel. 15, Created)                |   |                       |         |  |  |  |  |  |  |  |  |  |
| DT   | 01-OCT-2000   | (TrEMBLrel. 15, Last sequence update)   |   |                       |         |  |  |  |  |  |  |  |  |  |
| DT   | 01-JUN-2002   | (TrEMBLrel. 21, Last annotation update) |   |                       |         |  |  |  |  |  |  |  |  |  |
| DE   | 3'-phosphoadenosine 5'-phosphosulfate synthetase.                     |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| GN   | PAPS.   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| OS   | Homo sapiens (Human).   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| OC   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| OC   | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.            |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| OX   | NCBI_TaxID=9606;  |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| ON   | [1]   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RP   | SEQUENCE FROM N.A.  |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RC   | TISSUE=BRAIN;   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RX   | MEDLINE=20145453; PubMed=10679223;                                    |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RA   | Xu X.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,     |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RA   | Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,            |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RA   | Weinshilboum R.M.;  |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RT   | "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS1) and |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RT   | PAPS2: gene cloning, characterization and chromosomal localization."; |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| RL   | Biochem. Biophys. Res. Commun. 268:437-444,(2000).                    |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | EMBL:   | AF105227; AAF40236.1; ..                |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | InterPro:   | IPR002891; APS_Kinase.                  |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | InterPro:   | IPR002650; ATP-sulfurylase.             |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | Pfam:   | PF01583; APS_kinase; 1.                 |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | Pfam:   | PF01747; ATP-sulfurylase; 1.            |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | ProDom:   | PD002350; APS_kinase; 1.                |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | ProDom:   | PD002381; ATP-sulfurylase; 1.           |   |                       |         |  |  |  |  |  |  |  |  |  |
| DR   | TIGRFAMS:   | TIGR00455; apsk; 1.                     |   |                       |         |  |  |  |  |  |  |  |  |  |
| SQ   | SEQUENCE 624 AA; 70859 MW; 60DC9B943E7B75ED CRC64;                    |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| Query Match  |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| Best Local Similarity 79.48; Score 2596; DB 4; Length 624; |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |
| Matches 471; Conservative                                  |   |   |   |                       |         |  |  |  |  |  |  |  |  |  |



Query Match 68.3%; Score 2233.5; DB 5; Length 618;  
Best Local Similarity 68.2%; Pred. No. 2.5e-171;  
Matches 410; Conservative 75; Mismatches 115; Indels 1; Gaps 1;  
QY 10 ENQOKSTNVYQAHVSRNKRQGVGTGRCGCTVWLTGLSGAGKTTISFALEEYLVSH 69  
DB 14 EKPPMSNIVYQHHVSRNKRQGVGTGRCGCTVWLTGLSGAGKTTISFALEEYLVSH 72  
QY 70 AIPCYSLDGDNVHGLNRNLGSPGDREENIRRIAEVAKLFADAGLVLCITISFPAKDR 129  
DB 73 GIPAYSLDGNVHGLNRNLGSPGDREENIRRIAEVAKLFADAGLVLCITISFPAKDR 132  
QY 130 ENARKIHESAGLPFFEIFVDAPLNICESRDVKGILYKRRAGEIKGFTGIDSDEYKPTPE 189  
DB 133 QSARRVHNKSLPFFEIFVDTPLOVCEGRDVKGLYKRRAGEIKGFTGIDSDEYKPTPE 192  
QY 190 RVLKTNLSYSDCVHVVQLQEQNIIVPYTIIDHILFVPEKNDHVAEATLSLSI 249  
DB 193 VVTKTADMPVEDCMVYVVKELMERNIIPCETTLPVMEFTSHSOVGGKRAEALPSIDI 252  
QY 250 TKLDQWVQLSEGWATPLPGLFMEKEKEYLQVHMFDITLLDDGVINMSIPIVLPVSAEDKTR 309  
DB 253 TKLDQWVQLSEGWATPLPGLFMEKEKEYLQVHMFDITLLDDGVINMSIPIVLPVSAEDKTR 312  
QY 310 LEGCSFVLHAGRRVAILRDAEFYEHKREKRSRVGTTCTKHPHKKMYMESGDMVLVG 369  
DB 313 VENAETALYDQVKAHLKPEFYPHLKEERSQWGTGSKNGHPHKKMYMESGDMVLVG 372  
QY 370 DLQVLEKIRWNGDQVRLTLPLELKOKCKEMNADAVAFOLRNPVNHGALLMODTCRL 429  
DB 373 DIEVLRIWNGDQVRLTLPLELKOKCKEMNADAVAFOLRNPVNHGALLMODTCRL 432  
QY 430 LERKYHVPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLPKSTIVAIFFSPMLYAG 489  
DB 433 VERGFKNPVLHPLGGWTKDDVPLDWRMKOHAALVEEGVLPKSTIVAIFFSPMLYAG 492  
QY 490 PTEVQWCHRSMTAGANFYIVGRDPAGMPHPETKDYETHTGKVLSPAGLTSVEIIP 549  
DB 493 PTEVQWCHRSMTAGANFYIVGRDPAGMPHPETKDYETHTGKVLSPAGLTSVEIIP 552  
QY 550 FRVAAYNKAAMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAPKAMKVLVDYR 609  
DB 553 FRVAAYNKAAMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAPKAMKVLVDYR 612  
QY 610 S 610  
DB 613 S 613

RESULT 9  
Q961A8 PRELIMINARY; PRT; 629 AA.  
AC Q961A8; (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE LD25351P.  
GN PAPS OR CG8363.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY051174; AAK93148.1; -  
DR FlyBase; FBgn020389; Paps.  
DR InterPro; IPR002891; APS\_Kinase.  
DR InterPro; IPR002650; ATP-sulfonylase.  
DR Pfam; PF01593; APS\_kinase; 1.  
DR Pfam; PF01747; ATP-sulfonylase; 1.  
DR ProDom; PD002350; APS\_kinase; 1.  
DR ProDom; PD002381; ATP-sulfonylase; 1.  
DR TIGRFAMs; TIGR00455; aps; 1.  
SQ SEQUENCE 629 AA; 71146 MW; 6FF3F8797D98CCB3 CRC64;  
Query Match 62.3%; Score 2039; DB 5; Length 629;  
Best Local Similarity 63.1%; Pred. No. 1.2e-155;  
Matches 392; Conservative 95; Mismatches 120; Indels 14; Gaps 8;  
QY 2 SGIKKOTENQOKSTNVYQAHVSRNKRQGVGTGRCGCTVWLTGLSGAGKTTISFA 61  
DB 7 SSKKROKT-CLQVATNVTEQKHVTRTNKGNLGLCRGCTVWLTGLSGAGKTTISFA 65  
QY 62 LEBYLSHAIPOYSLDGNVHGLNRNLGSPGDREENIRRIAEVAKLFADAGLVLCITISF 121  
DB 66 LEAYLVSRGIPAYGLDGNIRTGLNGLGFTPADREENIRRVGEVAKLFADSGVVAICSF 125  
QY 122 ISFPAKDRENAKIHESAGLPFFEIFVDAPLNICESRDVKGILYKRRAGEIKGFTGIDS 181  
DB 126 VSPFADRENAKIHESAGLPFFEIFVDAPLNICESRDVKGILYKRRAGEIKGFTGIDS 185  
QY 182 YKPEPTEPVLKTNLSYSDCVHVVQLQEQNIIVPYTIIDHILFVPEKNDHVAEATLSLSI 237  
DB 186 YKPEPTEPVLKTNLSYSDCVHVVQLQEQNIIVPYTIIDHILFVPEKNDHVAEATLSLSI 244  
QY 238 RABAETLPSLSITKLDQWVQLSEGWATPLPGLFMEKEKEYLQVHMFDITLLDDGVINMSIPIVLPVSAEDKTR 293  
DB 245 RHEAESLQATEISITVELQWVLAEGWAYPLRGFMREDEYLTQTHFNTLQSGMDGYREN 304  
QY 294 MSTPIVLPVSAEDKTRLEGCSEKFLAHGRRVAILRDAEFYEHKREKRSRVGTTCTKHPHKKMYMESGDMVLVG 353  
DB 305 HSPVIVLSATQADKRLDCCSUTLKYQRAVAILRDAEFYEHKREKRSRVGTTCTKHPHKKMYMESGDMVLVG 364  
QY 354 PHIKMYMESGDMVLVGDLQVLEKIRWNGDQVRLTLPLELKOKCKEMNADAVAFOLRNPVNHGALLMODTCRL 413  
DB 365 PYSKQVYESDYLVGGDLAVIERIRWEDGLDQVRLTPNELRRRFKELNADAFQALRNP 424  
QY 414 VNHGALLMODTCRLLEKIRWNGDQVRLTLPLELKOKCKEMNADAVAFOLRNPVNHGALLMODTCRL 473  
DB 425 INHGHALLMODTCRLLEKIRWNGDQVRLTLPLELKOKCKEMNADAVAFOLRNPVNHGALLMODTCRL 484  
QY 474 KSTIVAIFFSPMLYAGTPEVQWCHRSMTAGANFYIVGRDPAGMPHPETKDYETHTGKVLSPAGLTSVEIIP 529  
DB 485 EDTVLAIFSPMAYAGTPEVQWCHRSMTAGANFYIVGRDPAGMPHPETKDYETHTGKVLSPAGLTSVEIIP 544  
QY 530 THGKVLSPAGLTSVEIIPFRVAAYNKAAMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAPKAMKVLVDYR 589  
DB 545 THGARVLKMAQGLDSMEILPFRVAAYDKSASRAFFEPKRDKEFEFISGTRMRKLAREGENPPDGFMAPKAMKVLVDYR 604  
QY 590 NPPDGFMAPKAMKVLVDYR 610  
DB 605 SPPDGFMEPEAWRILATYYQN 625

RESULT 10  
Q9VW48 PRELIMINARY; PRT; 630 AA.  
AC Q9VW48; O02638;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE Bifunctional ATP sulfurylase/APS kinase (EC 2.7.7.4) (Sulfate  
DE adenylyltransferase) (Sulfate adenylyl transferase) (ATP-sulfonylase)  
DE (PAPS protein) (CG8363 protein).  
GN PAPS OR CG8363.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

18 QVATNVTEQKHVHTRTRGNKGLCRGFRGCTVWLTLGSLGAGKTSIAPELAEYLVSRGIP 77  
 73 CYSLDGDNVRHGLNRLNLFSPGDREENIRIAEYAKLFADAGLVCITTSIFSPFAKDRENA 132  
 78 AYGLDGNIRTLGNLKNLGFTPADREENIRRVGEVAKLFADSGVVAICSFVSPFADDRESS 137  
 133 RKIHESAGLPFFEIFVDAPLNICESRDVKGLYKKARAGEIKGFTGIDSDYKPKETPERVL 192  
 138 RKTHKDAGLKFEYFVDTPLDVCETRDVKGLYKKAREGVIRKGTGITOQEYERPMQPELV 197  
 193 KTNLSIVSDCVHGVVQLLEQONTVPYTI IKDIH--ELFVPEV-KLDHVRAEAEFLPSLS 248  
 198 NTHGYTVRESTQKLVTLLLEGEIGIIPRS-LRDVDLLPELYPSESSTATEALRHAESLQAIE 256  
 249 ITKLDLQWQVVLSEGWATPLKGFMRKEEYELQVMHFDTLID--DGI--NMSIPIVLPVSA 304  
 257 ISVLEQWQVLAEGWAYPLRGPMREDEYLTQTLHENTLQSGMDGSGYRENHSHVPIVLSATQ 316  
 305 EDKTRLEGCKFVLANGRRVAILRDAEFYEHRKEERCVRVWGTTCTKHPHIKVMWESGD 364  
 317 ADKRLDGCSSLTKYQGVKAVAILRRPEFYFQRKEERLARQFTSNPNHPYSKQYVESGD 376  
 365 WLVGGLDVLLEKIRWNGDLOYRLTPLELKKCKCKEMNADAVFATQLRNPVINGHALLMOD 424  
 377 YLVGGDLAVTERIRWEDGDOYRUTPNELRRRFEKLNADAFATQLRNPVINGHALLMOD 436  
 425 TCRRLLEGRYKHPVLLLRPLGGWTKDDVDPLDMKMHAAVLEGVLDLPKSTIVAIFFSP 484  
 437 TRQLLEGRGFKQPVLLLRPLGGWTKDDVDPLDMKHOQAVLDAGVLRREDTVLAIFESP 496  
 485 MLYAGPTEVQWCHRSRMAGANFYIVGRDPAGMHPH--ETKKD--LYEPTHGKGVLSMAP 540  
 497 MMYAGPTEVQWHAHAKARNAGANFYIVGRDPAGMHPAKETYPDGNLYDATHGARVYLKMAQ 556  
 541 GLTSVETLIPRVAAYNKKAKMDFYDVPARHNEFDYFISCTRMRKLAREGENPPDGFMAPKA 600  
 557 GLDSMETLIPRVAAYDKSASRMAFEPKRDDEFESIGTKMRKTLAKTGASPPDGMEPEA 616  
 601 WKVLTIDYRS 610  
 617 WRILATYYQN 626  
 RESULT 11  
 Q22501 PRELIMINARY; PRT; 652 AA.  
 AC Q22501;  
 DT 01-NOV-1996 (TReMBUrel. 01, Created)  
 DT 01-NOV-1996 (TReMBUrel. 01, Last sequence update)  
 DE 01-JUN-2002 (TReMBUrel. 21, Last annotation update)  
 DE T14G10.1 protein.  
 GN T14G10.1  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wild A.;  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RL SEQUENCE FROM N.A.  
 RP MEDLINE=99069613; PubMed=9851916;  
 RX none;  
 RA "genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z6880; CAA93098.1; .  
 DR InterPro; IPR002891; APS\_kinase.  
 DR InterPro; IPR002650; ATP\_sulfurylase.  
 DR Pfam; PF01583; APS\_kinase; 1.  
 DR Pfam; PF01747; ATP\_sulfurylase; 1.  
 DR ProDom; PD002350; APS\_kinase; 1.

[illegible]

```

RESULT 12
Q9UIR2
ID Q9UIR2 PRELIMINARY; PRT; 265 AA.
AC Q9UIR2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE ATP sulfurylase/APS kinase isoform SK2 (EC 2.7.4) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98337975; PubMed=9671738;
RA Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr.,
RA Deyruz A., Schwartz N.B.;
RT "A member of a family of sulfate-activating enzymes causes murine
RT brachymorphism."
RL proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).

```

[illegible]



RP SEQUENCE FROM N.A.  
RA Cheuk R., Chen C.J., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RT "Arabidopsis cDNA clones";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,  
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C.C., Toriumi M.,  
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,  
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,  
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,  
RA Ecker J.R., Theologis A.;  
RT "Full length cDNA of gene F5N5.6/AT3g222890 (GI:11994718).";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,  
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,  
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,  
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,  
RA Davis R.W., Ecker J.R., Theologis A.;  
RT "Arabidopsis Open Reading Frame (ORF) Clones";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP001300; BAB03034.1; -  
DR EMBL: AY054169; AAL06830.1; -  
DR EMBL: AF370492; AAK43869.1; -  
DR EMBL: AY050869; AAK92806.1; -  
DR EMBL: AY064648; AAL47359.1; -  
DR EMBL: AY091207; AAM14146.1; -  
DR InterPro: IPR002650; ATP-sulphurylase.  
DR Pfam: PF01747; ATP-sulphurylase; 1.  
DR ProDom: PD002381; ATP-sulphurylase; 1.  
DR TIGRFAMs: TIGR00339; sopt; 1.  
KW Kinase.  
SQ SEQUENCE 463 AA; 51459 MW; F05DFD8F86CF25C8 CRC64;  
Query Match 39.8%; Score 1301; DB 10; Length 463;  
Best Local Similarity 62.7%; Pred. No. 2.4e-96;  
Matches 244; Conservative 60; Mismatches 81; Indels 4; Gaps 4;  
QY 226 ELFPENKLDHVRAEAEETLPSLITKDLQWQVLSGEGWATPLKGFMRKEKYLQVWHFDT 285  
DB 60 ELVEEPKRRKKKEAAADLPVELTAIDLQWHLVSEGWSPLGFGFMRSEFLQTLHFN 119

QY 286 L-LDGG-VINMSIPVLVPSAEDKTRLEGCSKFVLAHG-GRRVAILRDAEFYEHKRERC 342  
DB 120 LRLDDGSSVNNKSPVILVLAIDDEQKARIGESTRFVAFNSDGNPVALSDIEIYKHKPERI 179  
QY 343 SRVWGTTCTKPHI-KMYMESGDWLVGDLQVLEKIRWNGDLQVRLPLEKQKCKEMN 401  
DB 180 ARTWGTAPGLPYVDEAITHAGNWLIGDLVLEPKYVNGDLDRFLSPAELRKELEKRN 239  
QY 402 ADAVFAFQLRNPVNHGHALMODTCRRLLRGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 461  
DB 240 ADAVFAFQLRNPVNHGHALMODTCRRLLRGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 299  
QY 462 HAAVLEEGVLDPKSTIVAFITFSPMLYAGTVEYQWHCRCRSMIAGANFYIVGRDPAGMPHPE 521  
DB 300 HEKVLEDGVDPEITTVSIFPSPMHYAGTVEYQWHAHAKARINAGANFYIVGRDPAGMGHPV 359  
QY 522 TKKOLYETHGCKVLSMAPGLTSVEIIPFVAAYNKAKAMDFDPARHNEFDISGPRM 581  
DB 360 EKRDLYDADHGKVKVLSMAPGLERLNLPRVAAAYDKTQGMKAFDFPSRPQDFLISGTNM 419  
QY 582 RKLAREGENPPDGFMAPKAWKVLTDYYS 610  
DB 420 RTLAKNNENPPDGFMCPCGKVKLVDDYYS 448  
QY 424 AA.  
DB 424 AA.  
RESULT 14  
Q43170 PRELIMINARY; PRT; 424 AA.  
AC Q43170;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Sulfate adenylyltransferase (EC 2.7.7.4).  
GN STWET3-1.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DESIREE; TISSUE=LEAF;  
RX MEDLINE=95004649; PubMed=7920699;  
RA Klonus D., Hoefgen R., Willmitzer L., Riesmeier J.W.;  
RT "Isolation and characterization of two cDNAs encoding ATP-sulphurylase from potato by complementation of a yeast mutant.";  
RL Plant J. 6:105-112(1994).  
DR EMBL: X75041; CAA52953.1; -  
DR InterPro: IPR002650; ATP-sulphurylase.  
DR Pfam: PF01747; ATP-sulphurylase; 1.  
DR ProDom: PD002381; ATP-sulphurylase; 1.  
DR TIGRFAMs: TIGR00339; sopt; 1.  
KW Nucleotidyltransferase; Transferase.  
SQ SEQUENCE 424 AA; 47519 MW; E82A27DC11350ABC CRC64;  
Query Match 39.7%; Score 1299; DB 10; Length 424;  
Best Local Similarity 61.7%; Pred. No. 3e-96;  
Matches 240; Conservative 64; Mismatches 81; Indels 4; Gaps 4;  
QY 226 ELFPENKLDHVRAEAEETLPSLITKDLQWQVLSGEGWATPLKGFMRKEKYLQVWHFDT 285  
DB 21 ELFVDESQRDLKRAEALPKIKTKIDVWVHLSEGWSPLGFGFMRSEFLQTLHFN 80  
QY 286 L-LDGG-VINMSIPVLVPSAEDKTRLEGCSKFVLAHG-GRRVAILRDAEFYEHKRERC 342  
DB 81 TRLEDGSSVNNKSPVILVLAIDDSQKQIDGSSVALVVDGNPNPAILTDIEYKHKPERI 140  
QY 343 SRVWGTTCTKPHI-KMYMESGDWLVGDLQVLEKIRWNGDLQVRLPLEKQKCKEMN 401  
DB 141 ARTWGTAPGLPYVDEAITHAGNWLIGDLVLEPKYVNGDLDRFLSPAELRKELEKRN 200  
QY 402 ADAVFAFQLRNPVNHGHALMODTCRRLLRGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 461

```
|||||
201 ADAVAFQLRNPVNHGHALMTDTRRRLLJEMGYKNPVLLHPLGGYTKEDDVP LHRMKQ 260
462 HAAVLEEGVLDPKSTIVAFPPSPMLYAGPTEVQWCHRSRMAGANFYIVGRDPAGMPHPE 521
261 HEKVLGGVLDPTTVLSIFPPSPMHYAGTEVQWCHAKARINAGANFYIVGRDPAGMSHPL 320
522 TKKDLYEPTHGGKVLVSMAPGLTSVEIIPRVAAYNKKAKKAMDYDPAHNEFDIFSGTRM 581
321 EKRDLYDADHGKVKVLSMAPGLERLNLIPKVAAYDKTKNGMAFFDPSRQDFFISGTRM 380
582 RKLAREGENPPDGFMAPKANKVLTYYRS 610
381 RLAKTKESPDPDGFMCPPGKVKLVLEYYS 409
```

## RESULT 15

```
Q42520 PRELIMINARY; PRT; 465 AA.
AC Q42520;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP sulfurylase precursor (EC 2.7.7.4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96019964; PubMed=7487067;
RA Murillo M., Leustek T.;
RT "Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and
RT Escherichia coli are functionally equivalent but structurally and
RT kinetically divergent: nucleotide sequence of two adenosine-5'-
RT triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis
RT of a recombinant enzyme.";
RL Arch. Biochem. Biophys. 323:195-204 (1995).
DR EMBL; U06275; AAA92350.1; -
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002381; ATP-sulfurylase; 1.
DR TIGRFAMs; TIGR00339; sotp; 1.
KW Chloroplast; Nucleotidyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 50 CHLOROPLAST.
FT CHAIN 51 465 POTENTIAL.
SQ SEQUENCE 465 AA; 52029 MW; 4405F3E2AEF39735 CRC64;
```

```
Query Match 39.2%; Score 1282; DB 10; Length 465;
Best Local Similarity 62.3%; Pred. No. 8.le-95;
Matches 243; Conservative 60; Mismatches 81; Indels 6; Gaps 15;

QY 226 ELFVPEKLDHVRAEATLPSLSITKLDQVQLSEGWATPLKGFMRKEYLVQVHFDT 285
DB 62 DLVPEPRREKKHEADLPRLVLTALDQVHVLSEGWASPLRGFMRESEFLQTLHFNL 121
QY 286 L-LDQG-VINMSIPVLSAEDKTRLEGCSKFV--LAHGGRVAILRDAEYEHKKEER 341
DB 122 LNLDDGSVVNMSVPIVLAIDDOQA-LIGESKRVSLVSDSDNPIAILNDIEIYKHPKEER 180
QY 342 CSRWGTTCIKPHI-KWYMESGDWLVGDDLOVLEKIRWNDGLDQYRLTPLELKCKREM 400
DB 181 IARTWGTTPAGLPYVEEAITNAGDWLIGGDLEVPVYKNDGLRFLRSPFLRKELEKR 240
QY 401 NADAVAFQLRNPVNHGHALMODTCRLLRGYKHPVLLHPLGGWTKDDVPLDWRMK 460
DB 241 GADAVAFQLRNPVNHGHALMTDTRRLLEMGYKNPILLLHPLGGFTKADDVPLSWRMK 300
QY 461 QHAAVLEEGVLDPKSTIVAFPPSPMLYAGPTEVQWCHRSRMTAGANFYIVGRDPAGMPHP 520
DB 301 QHEKVLGGVLDPTTVLSIFPPSPMLYAGPTEVQWCHAKARINAGANFYIVGRDPAGMGHP 360
```

```
QY 521 ETKKDLYEPTHGGKVLVSMAPGLTSVEIIPRVAAYNKKAKKAMDYDPAHNEFDIFSGTR 580
DB 361 VEKRDLYDADHGKVKVLSMAPGLERLNLIPRVAAYDKTKNGMAFFDPSRQDFFISGTR 420
QY 581 MRKLAREGENPPDGFMAPKANKVLTYYRS 610
DB 421 MRALAKNRENPPDGFMCPPGKVKLVLEYYS 450
```

Search completed: May 30, 2003, 12:19:39

Job time : 91 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 14:58:56 ; Search time 4796 seconds  
(without alignments)  
11195.721 Million cell updates/sec

Title: US-09-898-165b-9  
Perfect score: 1845  
Sequence: 1 atgtcgggatacaagaagca.....ggtccctggagaagaactaa 1845

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Match | Length | DB | ID        | Description        |
|------------|--------|-------|--------|----|-----------|--------------------|
| 1          | 1845   | 100.0 | 2014   | 9  | AF091242  | AF091242 Homo sapi |
| 2          | 1841.8 | 99.8  | 2424   | 9  | BC009894  | BC009894 Homo sapi |
| 3          | 1840.2 | 99.7  | 1845   | 9  | AF313907  | AF313907 Homo sapi |
| 4          | 1838.6 | 99.7  | 3774   | 9  | AF074331  | AF074331 Homo sapi |
| 5          | 1816.8 | 98.5  | 1909   | 9  | AF173365  | AF173365 Homo sapi |
| 6          | 1813.6 | 98.3  | 2377   | 9  | AF150754  | AF150754 Homo sapi |
| 7          | 1483   | 80.4  | 3480   | 10 | AF251798  | AF251798 Cavia por |
| 8          | 1429.2 | 77.5  | 2000   | 10 | AF085144  | AF085144 Mus muscu |
| 9          | 1406.4 | 76.2  | 1996   | 10 | AF052453  | AF052453 Mus muscu |
| 10         | 1005   | 54.5  | 2479   | 10 | MM034883  | U34883 Mus musculu |
| 11         | 990    | 53.7  | 1875   | 6  | E27970    | E27970 APS kinase/ |
| 12         | 990    | 53.7  | 2265   | 9  | AF033026  | AF033026 Homo sapi |
| 13         | 990    | 53.7  | 2506   | 6  | AR044160  | AR044160 Sequence  |
| 14         | 990    | 53.7  | 2646   | 6  | AX281701  | AX281701 Sequence  |
| 15         | 988.6  | 53.6  | 2430   | 9  | BC011392  | BC011392 Homo sapi |
| 16         | 988.4  | 53.6  | 2537   | 9  | AF105227  | AF105227 Homo sapi |
| 17         | 986.8  | 53.5  | 2511   | 9  | HSPAPSSYN | Y10387 H.sapiens m |
| 18         | 983.6  | 53.3  | 1875   | 9  | AF016496  | AF016496 Homo sapi |
| 19         | 972.4  | 52.7  | 2282   | 9  | HSU53447  | U53447 Homo sapien |
| 20         | 943.2  | 51.1  | 2341   | 10 | AF004875  | AF004875 Cavia por |
| 21         | 828.4  | 44.9  | 2819   | 3  | UUNPASV   | L39001 Urechis cau |
| 22         | 734.6  | 39.8  | 2378   | 3  | AB036852  | AB036852 Ciona int |
| 23         | 639    | 34.6  | 2400   | 3  | AY051724  | AY051724 Drosophil |
| 24         | 620    | 33.6  | 2445   | 3  | DMPAPSSYN | Y12861 D.melanogas |
| 25         | 520.8  | 28.2  | 4407   | 2  | AC018153  | AC018153 Drosophil |
| 26         | 520.8  | 28.2  | 150884 | 2  | AC009382  | AC009382 Drosophil |
| 27         | 520.8  | 28.2  | 167816 | 3  | AC009381  | AC009381 Drosophil |
| 28         | 520.8  | 28.2  | 249805 | 3  | AE003515  | AE003515 Drosophil |
| 29         | 484.8  | 26.3  | 990    | 6  | E09951    | E09951 cDNA encodi |
| 30         | 391    | 21.2  | 1790   | 8  | AF016305  | AF016305 Zea mays  |
| 31         | 384    | 20.8  | 1662   | 8  | ATU40715  | U40715 Arabidopsis |
| 32         | 384    | 20.8  | 1680   | 8  | AY06276   | U06276 Arabidopsis |
| 33         | 384    | 20.8  | 1695   | 8  | AY086101  | AY086101 Arabidops |
| 34         | 384    | 20.8  | 1706   | 8  | ATMET31   | X79210 A.thaliana  |
| 35         | 383    | 20.8  | 1697   | 6  | ARI82690  | ARI82690 Sequence  |
| 36         | 382.4  | 20.7  | 1650   | 8  | AY093049  | AY093049 Arabidops |
| 37         | 376.8  | 20.4  | 1585   | 8  | AF403295  | AF403295 Allium ce |
| 38         | 373.6  | 20.2  | 1635   | 8  | AF212154  | AF212154 Allium ce |
| 39         | 371.8  | 20.2  | 1770   | 8  | STMET32   | X79053 S. tuberosu |
| 40         | 371    | 20.1  | 1715   | 8  | AF195511  | AF195511 Brassica  |
| 41         | 370    | 20.1  | 1749   | 8  | STATPS    | X75041 S.tuberosum |
| 42         | 368.2  | 20.0  | 1890   | 6  | ARI82691  | ARI82691 Sequence  |
| 43         | 368    | 19.9  | 2280   | 6  | ARI82689  | ARI82689 Sequence  |
| 44         | 365.6  | 19.8  | 1411   | 8  | AY064648  | AY064648 Arabidops |
| 45         | 365.6  | 19.8  | 1423   | 8  | AY091207  | AY091207 Arabidops |

ALIGNMENTS

RESULT 1  
AF091242  
LOCUS AF091242  
DEFINITION Homo sapiens ATP sulfurylase/APS kinase 2 mRNA, complete cds.  
ACCESSION AF091242  
VERSION AF091242.1 GI:3769609  
KEYWORDS Homo sapiens.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2014)  
AUTHORS ul Haque, M.F., King, L.M., Krakow, D., Cantor, R.M., Rusiniak, M.E.,  
Swank, R.T., Superti-Furga, A., Haque, S., Abbas, H., Ahmad, W.,  
Ahmad, M. and Cohn, D.H.

|                            |  |
|----------------------------|--|
| TITLE                      | Mutations in orthologous genes in human spondyloepimetaphyseal dysplasia and the brachymorphic mouse   |
| JOURNAL                    | Nat. Genet. 20 (2), 157-162 (1998)   |
| MEDLINE                    | 98372651   |
| PUBMED                     | 9771708  |
| REFERENCE                  | 2 (bases 1 to 2014)  |
| AUTHORS                    | ul Haque, M.F., King, L.M., Krakow, D., Cantor, R.M., Rusiniak, M.E., Swank, R.T., Supertli-Furga, A., Haque, S., Abbas, H., Ahmad, W., Ahmad, M. and Cohn, D.H.   |
| TITLE                      | Direct Submission  |
| JOURNAL                    | Submitted (08-SEP-1998) Pediatrics, Medical Genetics, Cedars-Sinai Research Institute, 8700 Beverly Blvd., Los Angeles, CA 90048, USA  |
| FEATURES                   | Location/Qualifiers  |
| source                     | 1..2014  |
|                            | /organism="Homo sapiens"   |
|                            | /db_xref="taxon:9606"  |
|                            | /chromosome="10"   |
|                            | /map="10q24"   |
|                            | /tissue_type="cartilage"   |
|                            | 81..1925   |
|                            | /note="ATPSK2"   |
|                            | /codon_start=1   |
|                            | /product="ATP sulfurylase/APS kinase 2"  |
|                            | /protein_id="AAC64583.1"   |
|                            | /db_xref="GI:3769610"  |
|                            | /translation="MSGIKKQKTNQKSTNVVYQAHVSRNKRQVVGTRGFRGCT VWTGSGAKTISFALEYLVSHAIPCYSLDQVNRHGLNRLNGSPDREINRR IAELVADAGLVCTISFISPAKPRENARKITHESAGLFFEIFVDAPLNIRESRDV KGLYKRAKEIGFTIDSDYKPEPTNRLNSTVSDCVHQVVELLQEQNIQVY TLIKDHELFPVNPDLHDVRAEATLPSLSITKLDQWVOVLSEGNATPLKGMREKE YIIVMHFDITLDGVNIMSPIVLPVSAEDKTRLEGSKFVLAHGRRVAILRDAPFY EHRKERCSRVWCTCTKHPH IKMVESGDMVLGSDLOVLEKTRWNGDLOYRLTPLE LKQCKEMRNADVAFQLRNPVINGHALLMQDRCRLLEKGYKHPVLLHLPGLGWTKD DVLPLWRMOKHAUVEEGVLDPKSTIVAFIPSPMLYAGPTEVOMHCRSRMIAGANFY IVRGDPAGMHPSTKDLNLEPTEHGKVLSPAGLTISVEIIFRVAAYNKAKKAMDFVD PARHNEFDFTSGTRMKLARENGENPPDGFMAKWKVLTIDYRSLEKN" |
|                            | 1504   |
| variation                  | /note="in SEMD Pakistani type"   |
|                            | /replace="a"   |
| BASE COUNT                 | 515 a 492 c 533 g 474 t  |
| ORIGIN                     |  |
| Query Match                | 100.0%; Score 1845; DB 9; Length 2014;   |
| Best Local Similarity      | 100.0%; Pred. No. 0;   |
| Matches 1845; Conservative | 0; Mismatches 0; Indels 0; Gaps 0;   |
| QY                         | 1 ATCTCGGGATCAAGAGCAAAAGACGGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 60  |
| Db                         | 81 ATCTCGGGATCAAGAGCAAAAGACGGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 140  |
| QY                         | 61 CAGGCCACCATCTGACAGCAATAAGAGAGGCAAGTGTGGAACAAGGGGTGGTTC 120  |
| Db                         | 141 CAGGCCACCATCTGACAGCAATAAGAGAGGCAAGTGTGGAACAAGGGGTGGTTC 200   |
| QY                         | 121 CGAGGATGATCCGTGGTGTAAACAGGTCTCTCTGGTGTGCTGGAACCAACGATAAGTTT 180  |
| Db                         | 201 CGAGGATGATCCGTGGTGTAAACAGGTCTCTCTGGTGTGCTGGAACCAACGATAAGTTT 260  |
| QY                         | 181 GCCTCGGAGGATGATCTGTCTCCCATGCCATCCCTTTGTTACTCCCTGGATGGGACAAT 240  |
| Db                         | 261 GCCTCGGAGGATGATCTGTCTCCCATGCCATCCCTTTGTTACTCCCTGGATGGGACAAT 320  |
| QY                         | 241 GTCCTGATGGCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAATATC 300   |
| Db                         | 321 GTCCTGATGGCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAATATC 380   |
| QY                         | 301 CGCCGATGCTGAGGTGGCTAAGCTGTTTGTGTGATGCTGGTGTGCTGCTGATTAACAGC 360  |
| Db                         | 381 CGCCGATGCTGAGGTGGCTAAGCTGTTTGTGTGATGCTGGTGTGCTGCTGATTAACAGC 440  |
| QY                         | 361 TTATTTCTCCATTCGAAGGATCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 420  |
| Db                         | 441 TTATTTCTCCATTCGAAGGATCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 500  |

|    |   |
|----|---|
| QY | 421 CTGCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACGTA 480    |
| Db | 501 CTGCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACGTA 560    |
| QY | 481 AAAGCCCTCTATAAAGAGCCAGAGCTGGGAGATTAAGAGATTTACAGGTATTGATCT 540     |
| Db | 561 AAAGCCCTCTATAAAGAGCCAGAGCTGGGAGATTAAGAGATTTACAGGTATTGATCT 620     |
| QY | 541 GATTATGAGAACTGAACTCCTGAGCTGTCTTAAACCAATTTTCCACAGTGA 600           |
| Db | 621 GATTATGAGAACTGAACTCCTGAGCTGTCTTAAACCAATTTTCCACAGTGA 680           |
| QY | 601 GACTGTGTCCACCACTAGTGAACCTTCTGCAAGAGCAGAACTTGTACCTCTATACTATA 660   |
| Db | 681 GACTGTGTCCACCACTAGTGAACCTTCTGCAAGAGCAGAACTTGTACCTCTATACTATA 740   |
| QY | 661 ATCAAGATATCCACGAACTCTTTGTCGGGAAACAACTTGACCACTCCGAGCTGAG 720       |
| Db | 741 ATCAAGATATCCACGAACTCTTTGTCGGGAAACAACTTGACCACTCCGAGCTGAG 800       |
| QY | 721 GCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGTCTTTG 780        |
| Db | 801 GCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGTCTTTG 860        |
| QY | 781 AGCGAAGCTGGGCACTCCCTCAAAGTTTTCATGCGGGAGAGGAGTACTTACAGTT 840       |
| Db | 861 AGCGAAGCTGGGCACTCCCTCAAAGTTTTCATGCGGGAGAGGAGTACTTACAGTT 920       |
| QY | 841 ATGCACCTTTCACACCTCTAGATGATGGCTGATCAACATGACATGCCATCCCTTGTACTG 900  |
| Db | 921 ATGCACCTTTCACACCTCTAGATGATGGCTGATCAACATGACATGCCATCCCTTGTACTG 980  |
| QY | 901 CCCGCTCTCTCAGAGGATAAGACACGGCTGGAAGGGTGCACAAAGTTTGTCTGGCACAT 960   |
| Db | 981 CCCGCTCTCTCAGAGGATAAGACACGGCTGGAAGGGTGCACAAAGTTTGTCTGGCACAT 1040  |
| QY | 961 GGTGGAGCGAGGTAGTCTATCTTACGAGCGCTGAATTTCTATGAACACAGAAAAGAGAA 1020  |
| Db | 1041 GGTGGAGCGAGGTAGTCTATCTTACGAGCGCTGAATTTCTATGAACACAGAAAAGAGAA 1100 |
| QY | 1021 CGCTGTTCCCGTGTGGGGACAAACATGTACAAACACCCCATATCAAAATGGTGTG 1080     |
| Db | 1101 CGCTGTTCCCGTGTGGGGACAAACATGTACAAACACCCCATATCAAAATGGTGTG 1160     |
| QY | 1081 GAAAGTGGGACCTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATAAGATGGAAT 1140 |
| Db | 1161 GAAAGTGGGACCTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATAAGATGGAAT 1220 |
| QY | 1141 GATGGGCTGACCAATACCGTCTGACACCTCTGAGAGCTCAACAGAAATGTAAAGAAATG 1200 |
| Db | 1221 GATGGGCTGACCAATACCGTCTGAGAGCTCAACAGAAATGTAAAGAAATG 1280          |
| QY | 1201 AATGCTGATCGGTGTTTGCATTCCAGTTCGCAATCTGTCACAAATGGCCATGCCCTG 1260   |
| Db | 1281 AATGCTGATCGGTGTTTGCATTCCAGTTCGCAATCTGTCACAAATGGCCATGCCCTG 1340   |
| QY | 1261 TTGATGAGGACACCTGCGCAGGCTCTTAGAGAGGGGCTACAAAGCACCCTGCTCCTA 1320   |
| Db | 1341 TTGATGAGGACACCTGCGCAGGCTCTTAGAGAGGGGCTACAAAGCACCCTGCTCCTA 1400   |
| QY | 1321 CTACACCTCTGGCGGCTGACCAAGGATGACGATGCTCTAGACTGGCGATGAAG 1380       |
| Db | 1401 CTACACCTCTGGCGGCTGACCAAGGATGACGATGCTCTAGACTGGCGATGAAG 1460       |
| QY | 1381 CAGCAGCGGCTGTGCTCGAGGAGGGTCTCTGGATCCCAAGTCAACCAATTTGTGCCATC 1440 |
| Db | 1461 CAGCAGCGGCTGTGCTCGAGGAGGGTCTCTGGATCCCAAGTCAACCAATTTGTGCCATC 1520 |
| QY | 1441 TTTCCGCTCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCAGTCCAGTCCCGG 1500 |
| Db | 1521 TTTCCGCTCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCAGTCCAGTCCCGG 1580 |

```

QY 1501 ATGATTGGGGTGCACATTTCTACATTTGGGGAGGAGCCCTGCGAGGAATGCCCATCCT 1560
Db 1581 ATGATTGGGGTGCACATTTCTACATTTGGGGAGGAGCCCTGCGAGGAATGCCCATCCT 1640
QY 1561 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGGCTTGGAGCATGCCCT 1620
Db 1641 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGGCTTGGAGCATGCCCT 1700
QY 1621 GGCCTCACTCTGTTGGAATCATTCCATTCGAGTGGCTGCTACAAAGGCAAAAAA 1680
Db 1701 GGCCTCACTCTGTTGGAATCATTCCATTCGAGTGGCTGCTACAAAGGCAAAAAA 1760
QY 1681 GGCATGACTTCTATGATCCAGCAAGGACAAATGATTTGATCTCAGGAACCTGA 1740
Db 1761 GGCATGACTTCTATGATCCAGCAAGGACAAATGATTTGATCTCAGGAACCTGA 1820
QY 1741 ATGAGGAAGCTCGCCCGGAAGGAGAGAAATCCCGAGATGGCTTCATGGCCCCCAAGCA 1800
Db 1821 ATGAGGAAGCTCGCCCGGAAGGAGAGAAATCCCGAGATGGCTTCATGGCCCCCAAGCA 1880
QY 1801 TGGAGGTCCTGACAGATTTATTACAGGTCCTCGGAGAGAACTAA 1845
Db 1881 TGGAGGTCCTGACAGATTTATTACAGGTCCTCGGAGAGAACTAA 1925

```

```

RESULT 2
BC009894 2424 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:1916 IMAGE:2988345, mRNA, complete cds.
DEFINITION BC009894
ACCESSION BC009894
VERSION BC009894.1 GI:14602765
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2424)
Strausberg,R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabp-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fabey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 3 Row: p Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3769609.

Location/Qualifiers

```

1. .2424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:1916 IMAGE:2988345"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_15"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
121. .1965
/codon_start=1
/product="unknown (protein for MGC:1916)"
/protein_id="AAH09894.1"

```

FEATURES source

CDS

```

/db_xref="GI:14602766"
/translation="MSGIKKQKNTNYYQAHVHVRNKRKQVVGTRGFRGCT
VWLTGLSGAGTITISFALELYIVSHAIPIYSLDGNVHRLNRLNLSFGSPGREENIR
IAEVALFADAGLVCTISFISPAKDNENARKIHESAGLPEFEIVDCEKLVNEDV
KGLVKRARGIKGFTGDSYKPEKPELVKLTNLSTVDCVHOVVELLOEONIVPY
TLIKDIHELFPENKLDHVRAEETLPSLITKLDHVOVLSGEMATPKGPMKEKE
YLVQWFDITLDDGGVYNNSTIFIVLPMESADKRLGSKFVYLAHRRVRLDQRLPFE
EHRKEERSRVMTGTTCTKHPHIMVNESDMLVGGDLQVLEKIRWNGDLOQTRUPLF
LKQCKKENNADAVAFQRLNPNHNGHALLMQDTRRLRLRGYKHPVLHLLHMGWTK
DDVPLDMRKQHAALVEBGLDPKSTIVAFISPMLYAGTEVQWHCRSMILGAGNFY
IVGRDPAPHPETKKDLYEPTHGKVLSPAGLTVEIIPRVAAYNKKAKMDFDYD
PARNEPFIISGTRMRKLAREGENPPDGFMAPKWKVLTVDYVRSLEKN"

BASE COUNT      661 a  565 c  596 g  602 t
ORIGIN
Query Match      99.8%; Score 1841.8; DB 9; Length 2424;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1843; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTCGGGGATCAAGAAGCAAAAGACGAGAGAACCAAGCAAAATCCACCAATAGTCTAT 60
Db 121 ATGTCGGGGATCAAGAAGCAAAAGACGAGAGAACCAAGCAAAATCCACCAATAGTCTAT 180
QY 61 CAGCCCCACCATGTGAGCAGGAATAAGAGAGGCGCAAGTGGTGGACAAAGGGTGCGTTC 120
Db 181 CAGCCCCACCATGTGAGCAGGAATAAGAGAGGCGCAAGTGGTGGACAAAGGGTGCGTTC 240
QY 121 CGAGGATGTACCGTGTGCTAACAGGTCTCTCGTGTCTGGAAAAACACGATAAGTTTT 180
Db 241 CGAGGATGTACCGTGTGCTAACAGGTCTCTCGTGTCTGGAAAAACACGATAAGTTTT 300
QY 181 GCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 240
Db 301 GCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 360
QY 241 GTCCCTCATGGCTTAAACAGAAATCTCGGATTCCTCTCTGGGAGAGAGAGAAATATC 300
Db 361 GTCCCTCATGGCTTAAACAGAAATCTCGGATTCCTCTCTGGGAGAGAGAGAAATATC 420
QY 301 CCGCGGATGTCTGAGTGGCTAAGCTGTTTGTGTGATGTCTGTCTGTCTGATACACG 360
Db 421 CCGCGGATGTCTGAGTGGCTAAGCTGTTTGTGTGATGTCTGTCTGTCTGATACACG 480
QY 361 TTTATTTCTCCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATCAATCAGCAGG 420
Db 481 TTTATTTCTCCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATCAATCAGCAGG 540
QY 421 CTGCCATTTCTTGAATAATTTTGTAGATGCACCTCTAAATATTTGTGAAAAGCAGAG 480
Db 541 CTGCCATTTCTTGAATAATTTTGTAGATGCACCTCTAAATATTTGTGAAAAGCAGAG 600
QY 481 AAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTTAAAGGATTTACAGGTATTGATCT 540
Db 601 AAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTTAAAGGATTTACAGGTATTGATCT 560
QY 541 GATTATGAGAAACCTGAAACTCCTGAGCGTGTCTGTAAACCAATTTGTCCAGTGAGT 600
Db 661 GATTATGAGAAACCTGAAACTCCTGAGCGTGTCTGTAAACCAATTTGTCCAGTGAGT 720
QY 601 GACTGTGCCACAGGTACTGGAACCTTCTGCAAGAGCAGCAACATCTTACCCCTATCTATA 660
Db 721 GACTGTGCCACAGGTACTGGAACCTTCTGCAAGAGCAGCAACATCTTACCCCTATCTATA 780
QY 661 ATCAAGATATATCAGGAACCTTTTGTGCCGGAACCAACTTGACCACCTCCGAGCTGAG 720
Db 781 ATCAAGATATATCAGGAACCTTTTGTGCCGGAACCAACTTGACCACCTCCGAGCTGAG 840
QY 721 GCTGAACCTCTCCCTTCATTAATTAATTAAGTGGATCTCCAGTGGGTCCAGTTTGG 780
Db 841 GCTGAACCTCTCCCTTCATTAATTAATTAAGTGGATCTCCAGTGGGTCCAGTTTGG 900
QY 781 ACAGGAGCTGGGCCACTCCCTCAAGGTTTTCATCGGGGAGAGAGTACTTACAGGTT 840

```

Db 901 AGCGAAGCTGGGCGCACTCCCTCAAGGTTTCATCGGGGAGAGGAGTACTTACAGGTT 960  
QY 841 ATGCACATTTGACACCTCTGCTAGATGATGGGTGATCAACATGACATGCCCATTTGACTG 900  
Db 961 ATGCACATTTGACACCTCTGCTAGATGATGGGTGATCAACATGACATGCCCATTTGACTG 1020  
QY 901 CCCGTCTCTCAGAGNATAGACACGCTCGAAGGTGACAGAAAGTTTCTTGGGCACAT 960  
Db 1021 CCCGTCTCTCAGAGNATAGACACGCTCGAAGGTGACAGAAAGTTTCTTGGGCACAT 1080  
QY 961 GGTGGAGGAGGTAGTCTTACGAGACGCTGCAATCTTATGAACACACAAAGAGGAA 1020  
Db 1081 GGTGGAGGAGGTAGTCTTACGAGACGCTGCAATCTTATGAACACACAAAGAGGAA 1140  
QY 1021 CGCTGTCTCCGTGTTGGGGGACACATGTACAAACACCCCATATCAAAATGGTGATG 1080  
Db 1141 CGCTGTCTCCGTGTTGGGGGACACATGTACAAACACCCCATATCAAAATGGTGATG 1200  
QY 1081 GAAAGTGGGACATGGCTGGTGTGGAGACCTTCAGTGTCTGGAGAAATAAGATGAAT 1140  
Db 1201 GAAAGTGGGACATGGCTGGTGTGGAGACCTTCAGTGTCTGGAGAAATAAGATGAAT 1260  
QY 1141 GATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1200  
Db 1261 GATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1320  
QY 1201 AATGCTGATGCGGTGTTGCAATCCAGTTCGCAATCTCTGCAATGCGGCTGATGCCCTG 1380  
Db 1321 AATGCTGATGCGGTGTTGCAATCCAGTTCGCAATCTCTGCAATGCGGCTGATGCCCTG 1380  
QY 1261 TTGATGAGGACATGCGGACGCTCTAGAGAGGGGTACAGACACCGGCTCTCTCTA 1320  
Db 1381 TTGATGAGGACATGCGGACGCTCTAGAGAGGGGTACAGACACCGGCTCTCTCTA 1440  
QY 1321 CTACACCTCTGGGCGCTGACCAAGGATGACATGTGCTCTAGACTGGCGGATGAAG 1380  
Db 1441 CTACACCTCTGGGCGCTGACCAAGGATGACATGTGCTCTAGACTGGCGGATGAAG 1500  
QY 1381 CAGCAGCGGCTGTGCTGAGAGAGGGTCTCGATCCCAAGTCAACCAATGTTGGCATC 1440  
Db 1501 CAGCAGCGGCTGTGCTGAGAGAGGGTCTCGATCCCAAGTCAACCAATGTTGGCATC 1560  
QY 1441 TTTCCGTCTCCCATGTTATGCTGCCCCACAGAGTCCAGTGGGACGCTGAGTCCCGG 1500  
Db 1561 TTTCCGTCTCCCATGTTATGCTGCCCCACAGAGTCCAGTGGGACGCTGAGTCCCGG 1620  
QY 1501 ATGATTGCGGGTCCCAATTTCTACATTTGGGAGGAGCCCTGCAGGAATGCCCATCCT 1560  
Db 1621 ATGATTGCGGGTCCCAATTTCTACATTTGGGAGGAGCCCTGCAGGAATGCCCATCCT 1680  
QY 1561 GAAACCAAGAGATCTGATGAACCCACTATCGGGGCAAGTCTTGAGCATGCGCCCT 1620  
Db 1681 GAAACCAAGAGATCTGATGAACCCACTATCGGGGCAAGTCTTGAGCATGCGCCCT 1740  
QY 1621 GGCTCACCTCTGTGAAATCATTTCCATTCGAGTGGCTGCCCTACACAAAGCCAAAAA 1680  
Db 1741 GGCTCACCTCTGTGAAATCATTTCCATTCGAGTGGCTGCCCTACACAAAGCCAAAAA 1800  
QY 1681 GCCATGCACTCTATGATCCAGAGGCAATGAGTTTGAATTCATCTCAGGAATCGA 1740  
Db 1801 GCCATGCACTCTATGATCCAGAGGCAATGAGTTTGAATTCATCTCAGGAATCGA 1860  
QY 1741 ATGAGGAGCTGCCCGGAGAGAGATTCCTCCAGATGGCTTCATGCCCCCAAGCA 1800  
Db 1861 ATGAGGAGCTGCCCGGAGAGAGATTCCTCCAGATGGCTTCATGCCCCCAAGCA 1920  
QY 1801 TGGAGGCTGACATATTACAGTCCCTGGAGAAAGTAA 1845  
Db 1921 TGGAGGCTGACATATTACAGTCCCTGGAGAAAGTAA 1965

RESULT 3  
AF313907

LOCUS AF313907 1845 bp mRNA linear PRI 01-FEB-2001  
DEFINITION Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthase 2 alpha  
mRNA, complete cds.  
ACCESSION AF313907  
VERSION AF313907.1 GI:12642583  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1845)  
AUTHORS Fuda,H., Shimizu,C. and Strott,C.A.  
TITLE Human Bifunctional 3'-Phosphoadenosine 5'-phosphosulfate Synthase:  
Differential Expression of Isoforms and Effect of Polymorphisms on  
Activity  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1845)  
AUTHORS Shimizu,C., Fuda,H. and Strott,C.A.  
TITLE Direct Submission  
JOURNAL Submitted (17-OCT-2000) Section on Steroid Regulation, NICHD/NIH,  
9000 Rockville Pike, Bethesda, MD 20892-4510, USA

FEATURES  
source Location/Qualifiers  
1..1845  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

cds  
1..1845  
/function="catalyzes production of 3'-phosphoadenosine  
5'-phosphosulfate, the universal sulfonate donor molecule,  
from inorganic sulfate and ATP"

/codon\_start=1  
/product="3'-phosphoadenosine 5'-phosphosulfate synthase 2  
alpha"  
/protein\_id="AAK00296.1"  
/db\_xref="GI:12642584"

/translation="MSGIKKQKTEENQKSTNVVQAHVHNRKRGVVGTRGGPRGCT  
VWLTGLSGAGKTTISPALEYLYSHAIPICSLDGDNVRHGLNRLNGSPGDRNENRR  
IADVKLFADAGLVCTISFPAKRENAARKIHESAGLFEIFVDAPLNCEISRDV  
KGLYKRAAGEIKGTGIDSDYKPEPTELKLNLSVDCVHVQVVELLOENIVBY  
TITKDHELFVPEKLDHVAEATLPSLITLDLQVQLVSEGWAFLPKGMREKE  
YLVQHFDTLLDDGVINMSPIVLPVSAEDTKLEGSKVLAHGRVAILLRDAEY  
EHKREKRCRWGVTCTTCKPHIKMVGSGDVLGGDLQVLEKIRWNGDLQYRLTPE  
LKQCKERNADAVAFQLRNPVHALLMQDTRRRLLRGYKHPVLLHLLPLGGWTKD  
DVLPLWRMKQHAALVEGLDPKSTIVAFPSPLMAYAGTTEVQWCHRSRMAGANPY  
IVGRDPAGMHPETKDKLDEPTGGKGLSNAGLTISVEIIPFRVAATNKAAMDFYD  
PARHNEFDFTSGTRMKLAREGENPDGEMAPKAWKVLTDYRSLEKN"

BASE COUNT 494 a 436 c 496 g 419 t

Query Match 99.7%; Score 1840.2; DB 9; Length 1845;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1842; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCTCGGGGATCAAGAGCAAAAAGACGAGCAACAGCAAGAAATCCACCAATGTAGTCTAT 60  
Db 1 ATCTCGGGGATCAAGAGCAAAAAGACGAGCAACAGCAAGAAATCCACCAATGTAGTCTAT 60  
QY 61 CAGGCCCAACCATCTGAGCAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGGTTTC 120  
Db 61 CAGGCCCAACCATCTGAGCAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGGTTTC 120  
QY 121 CGAGGATGTACCGTGGCTTAACAGGTCTCTCTGGTGTCTGGAACAAACAGATAAGTTTT 180  
Db 121 CGAGGATGTACCGTGGCTTAACAGGTCTCTCTGGTGTCTGGAACAAACAGATAAGTTTT 180  
QY 181 GCCTCTGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTTACCTCCCTGGATGGGACAAT 240  
Db 181 GCCTCTGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTTACCTCCCTGGATGGGACAAT 240  
QY 241 GTCCGTCATGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGGAAAAATATC 300  
Db 241 GTCCGTCATGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGGAAAAATATC 300  
QY 301 CGCCGGATGTGCTGAGGTGGCTAAGCTGTGTGCTGATGCTGGTCTGCATTACCAACCAGC 360

Db 301 CGCCGAGTCTGAGTGGTAAAGCTTTGCTGATGCTGGTCTGCATACACG 360  
Qy 361 TTTATTTCTCATTCGCAAGGATCGTGAAGTCCCGCAAAATACATGAATACAGCAGG 420  
Db 361 TTTATTTCTCATTCGCAAGGATCGTGAAGTCCCGCAAAATACATGAATACAGCAGG 420  
Qy 421 CTGCAATCTTTGAAATATTGTAGATGCAACCTCTAAATATTGTGAAAGCAGAGCGTA 480  
Db 421 CTGCAATCTTTGAAATATTGTAGATGCAACCTCTAAATATTGTGAAAGCAGAGCGTA 480  
Qy 481 AAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATCT 540  
Db 481 AAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATCT 540  
Qy 541 GATTATGAGAACTGAAACTCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAGT 600  
Db 541 GATTATGAGAACTGAAACTCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAGT 600  
Qy 601 GACTGTCTCCACAGGTAGTGAACCTCTGCAAGAGCAGAACATTTGACCCCTATACATA 660  
Db 601 GACTGTCTCCACAGGTAGTGAACCTCTGCAAGAGCAGAACATTTGACCCCTATACATA 660  
Qy 661 ATCAAGATATCCACAGTCTTTGTCGGGAAACAACTTGACACAGTCCGAGCTGAG 720  
Db 661 ATCAAGATATCCACAGTCTTTGTCGGGAAACAACTTGACACAGTCCGAGCTGAG 720  
Qy 721 GCTGAAACTCTCCCTTCAATATCAATTAAGCTGATCTCCAGTGGTCCAGGTTTG 780  
Db 721 GCTGAAACTCTCCCTTCAATATCAATTAAGCTGATCTCCAGTGGTCCAGGTTTG 780  
Qy 781 AGCAAGGCTGGGCACTCCCTCAAGGTTTTCATCGGGAGAGAGTACTTACAGGTT 840  
Db 781 AGCAAGGCTGGGCACTCCCTCAAGGTTTTCATCGGGAGAGAGTACTTACAGGTT 840  
Qy 841 ATGCACTTTCACACCTCTAGATGATGGGTGATCAACATGACCATCCCATTTGACTG 900  
Db 841 ATGCACTTTCACACCTCTAGATGATGGGTGATCAACATGACCATCCCATTTGACTG 900  
Qy 901 CCCGCTCTCAGAGGATAAGACACGCTGGAAGGTCAGCAAGTGTCTGCTGGCACAT 960  
Db 901 CCCGCTCTCAGAGGATAAGACACGCTGGAAGGTCAGCAAGTGTCTGCTGGCACAT 960  
Qy 961 GTGGAGGAGGAGTACTTCTACAGAGCTGTAATCTATGAACACAGAAAGAGAA 1020  
Db 961 GTGGAGGAGGAGTACTTCTACAGAGCTGTAATCTATGAACACAGAAAGAGAA 1020  
Qy 1021 CGCTGTTCCGCTGTTTGGGGACAACTGTACAAACACCCCATATCAAAATGCTGAT 1080  
Db 1021 CGCTGTTCCGCTGTTTGGGGACAACTGTACAAACACCCCATATCAAAATGCTGAT 1080  
Qy 1081 GAAAGTGGGACTGGCTGTTGTTGGAGACCTTCAGGTGCTGGAGAAATAAGATGCAAT 1140  
Db 1081 GGAAGTGGGACTGGCTGTTGTTGGAGACCTTCAGGTGCTGGAGAAATAAGATGCAAT 1140  
Qy 1141 GATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1200  
Db 1141 GATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1200  
Qy 1201 AATGCTGATGGGTTGTTGATTCACCTGCGCAATCTGTCACATGCGCCCTG 1260  
Db 1201 AATGCTGATGGGTTGTTGATTCACCTGCGCAATCTGTCACATGCGCCCTG 1260  
Qy 1261 TTGATGAGGACACCTCGCAGGCTCTAGAGAGGCTTACAGACCCCGTCTCCTA 1320  
Db 1261 TTGATGAGGACACCTCGCAGGCTCTAGAGAGGCTTACAGACCCCGTCTCCTA 1320  
Qy 1321 CTACACCTCTGGGCGCTGGACCAAGGATGACGATGCTCTAGACTGCGGATGAAG 1380  
Db 1321 CTACACCTCTGGGCGCTGGACCAAGGATGACGATGCTCTAGACTGCGGATGAAG 1380  
Qy 1381 CAGCAGCGGCTGTGCTGAGGAGGCTGCTGATCCCAAGTCAACCATTTGTCATC 1440

Db 1381 CAGCAGCGGCTGTGCTGAGGAGGCTCTGATCCCAAGTCAACCATTTGTCATC 1440  
Qy 1441 TTTCCGCTCCCATTTATATGCTGGCCACACAGAGTCCAGTGCAGTGCAGTCCCGG 1500  
Db 1441 TTTCCGCTCCCATTTATATGCTGGCCACACAGAGTCCAGTGCAGTGCAGTCCCGG 1500  
Qy 1501 ATGATTGGGCTGCCAATTTCTACATTTGTGGGAGGACCTGCGAGGAATGCCCATCT 1560  
Db 1501 ATGATTGGGCTGCCAATTTCTACATTTGTGGGAGGACCTGCGAGGAATGCCCATCT 1560  
Qy 1561 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGTCTTCAGCATGGCCCT 1620  
Db 1561 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGTCTTCAGCATGGCCCT 1620  
Qy 1621 GGCCTCACTCTGTGGAATCATTTCCATTCCGAGTGGCTGCTACACAAAGCCAAAAA 1680  
Db 1621 GGCCTCACTCTGTGGAATCATTTCCATTCCGAGTGGCTGCTACACAAAGCCAAAAA 1680  
Qy 1681 GCCATGACTTCTATGATCCAGAGGACCAATGAGTTTGAATCTCAGGAATCTCGA 1740  
Db 1681 GCCATGACTTCTATGATCCAGAGGACCAATGAGTTTGAATCTCAGGAATCTCGA 1740  
Qy 1741 ATGAGGAGTCCGCGGAGGAGAGAAATCCCGAGATGGCTTCATGGCCCGCAAGCA 1800  
Db 1741 ATGAGGAGTCCGCGGAGGAGAGAAATCCCGAGATGGCTTCATGGCCCGCAAGCA 1800  
Qy 1801 TGGAGGTCCTGACAGATTTATACAGTCTCTGGAGAGAACTAA 1845  
Db 1801 TGGAGGTCCTGACAGATTTATACAGTCTCTGGAGAGAACTAA 1845

RESULT 4  
AF074331  
LOCUS  
DEFINITION Homo sapiens PAPS synthetase-2 (PAPSS2) mRNA, complete cds.  
ACCESSION AF074331  
VERSION AF074331.1 GI:5052074  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3774)  
AUTHORS Franzon, V.L., Gibson, M.A., Hatzinikolas, G., Cleary, E.G., Woolatt, E.  
and Sutherland, G.R.  
TITLE Direct Submission  
JOURNAL Submitted (24-JUN-1998) Pathology, University of Adelaide,  
Adelaide, SA 5005, Australia  
FEATURES  
Location/Qualifiers  
1..3774  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q23-24.1"  
1..3774  
/gene="PAPSS2"  
64..1908  
/gene="PAPSS2"  
/codon\_start=1  
/product="PAPS synthetase-2"  
/protein\_id="AAD38423.1"  
/db\_xref="GI:5052075"

gene  
CDS

|            |  |  |      |
|------------|--|--|------|
| Db         | 1024   | GGTGGACGGAGGTAGCTATCTTTACGAGACGCTGGAATTTCTATGAACACAGAAAGAGGAA  | 1081 |
| Qy         | 1021   | CGCTGTTCCCGCTGTTGGGGGACAACATGTACAAAACACCCCCCATATCAAAATGGTGTATG | 1080 |
| Db         | 1084   | CGCTGTTCCCGCTGTTGGGGGACAACATGTACAAAACACCCCCCATATCAAAATGGTGTATG | 1143 |
| Qy         | 1081   | GAAGTGGGACCTGGCTGTTGGTGGAGACCTTTACAGTGTCTGGAGAAAATAGATGGAAT    | 1140 |
| Db         | 1144   | GAAGTGGGACCTGGCTGTTGGTGGAGACCTTTACAGTGTCTGGAGAAAATAGATGGAAT    | 1203 |
| Qy         | 1141   | GATGGGCTGACCAATACCGTCTGACACCTCTCGAGCTCAAAACAGAAAATGTAAGAATAATG | 1200 |
| Db         | 1204   | GATGGGCTGACCAATACCGTCTGACACCTCTCGAGCTCAAAACAGAAAATGTAAGAATAATG | 1263 |
| Qy         | 1201   | AATGCTGATGCGGTGTTGGCATTTCCAGTTGCGCAATCTCTGCCAATGGCCATGCCCTG    | 1260 |
| Db         | 1264   | AATGCTGATGCGGTGTTGGCATTTCCAGTTGCGCAATCTCTGCCAATGGCCATGCCCTG    | 1323 |
| Qy         | 1261   | TTGATCGAGACACTGCGCGAGGCTCTAGAGAGGGCTACAAGCACCCTGCTCCCTA        | 1320 |
| Db         | 1324   | TTGATCGAGACACTGCGCGAGGCTCTAGAGAGGGCTACAAGCACCCTGCTCCCTA        | 1383 |
| Qy         | 1321   | CTACACCCCTGCGCGGCTGGACCAAGATGACGATGTGCCTCTAGACTGGCGGATGAAG     | 1380 |
| Db         | 1384   | CTACACCCCTGCGCGGCTGGACCAAGATGACGATGTGCCTCTAGACTGGCGGATGAAG     | 1443 |
| Qy         | 1381   | CAGCAGCGGCTGTGCTCGAGGAAGGGTCTGGATCCCAAGTCACACCATTTGTCGCATC     | 1440 |
| Db         | 1444   | CAGCAGCGGCTGTGCTCGAGGAAGGGTCTGGATCCCAAGTCACACCATTTGTCGCATC     | 1503 |
| Qy         | 1441   | TTTCCGCTCTCCCATGTTATGCTGGCCCCACAGAGGTCCAGTGGCACCTGCAGGTCCCGG   | 1500 |
| Db         | 1504   | TTTCCGCTCTCCCATGTTATGCTGGCCCCACAGAGGTCCAGTGGCACCTGCAGGTCCCGG   | 1563 |
| Qy         | 1501   | ATGATTGGGGTGCCAAATTTCTACATTTGGGAGAGGACCTTCGAGGAATGCCCATCCT     | 1560 |
| Db         | 1564   | ATGATTGGGGTGCCAAATTTCTACATTTGGGAGAGGACCTTCGAGGAATGCCCATCCT     | 1623 |
| Qy         | 1561   | GAACCAAGAAGGATCTGTATGAACCCACTCATTTGGGGGCAAGGCTTTGAGCATGGGCCCT  | 1620 |
| Db         | 1624   | GAACCAAGAAGGATCTGTATGAACCCACTCATTTGGGGGCAAGGCTTTGAGCATGGGCCCT  | 1683 |
| Qy         | 1621   | GGCCTCACCTCTGTGAAATCATTTCCAGTGGCTGGCTGCCTTACAAACAGCAAAAAA      | 1680 |
| Db         | 1684   | GGCCTCACCTCTGTGAAATCATTTCCAGTGGCTGGCTGCCTTACAAACAGCAAAAAA      | 1743 |
| Qy         | 1681   | GCCATGGACTTCATGATCCAGCAAGGCACATGAGTTTGACTTCATCTCAGGAACCTGA     | 1740 |
| Db         | 1744   | GCCATGGACTTCATGATCCAGCAAGGCACATGAGTTTGACTTCATCTCAGGAACCTGA     | 1803 |
| Qy         | 1741   | ATGAGGAAGCTGCCCGGGAAGAGAGATCCCCAGATGGCTTTCATGCCCCCAAGCA        | 1800 |
| Db         | 1804   | ATGAGGAAGCTGCCCGGGAAGAGAGATCCCCAGATGGCTTTCATGCCCCCAAGCA        | 1863 |
| Qy         | 1801   | TGGAAGGCTCTGACGATTTATACAGTCCCTGGAGAAAGCACTAA                   | 1845 |
| Db         | 1864   | TGGAAGGCTCTGACGATTTATACAGTCCCTGGAGAAAGCACTAA                   | 1908 |
| RESULT 5   |  |  |      |
| AF173365   |  |  |      |
| LOCUS      | AF173365 1909 bp mRNA linear PRI 23-NOV-1999   |  |      |
| DEFINITION | Homo sapiens ATP sulfurylase/AFPS kinase isoform SK2 mRNA, alternatively spliced, complete cds.                              |  |      |
| ACCESSION  | AF173365   |  |      |
| VERSION    | AF173365.1   |  |      |
| KEYWORDS   | GI:6466025   |  |      |
| SOURCE     | Homo sapiens.  |  |      |
| ORGANISM   | Homo sapiens   |  |      |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |  |      |
| AUTHORS    | Kurima, K., Waxman, M.L., Krishnan, S., Domowicz, M., Krueger, R.C., Jr.,  |  |      |





Db 1429 GACTGGCGGATGAGCAGCAGCGGCTGTCTCGAGGAAGGGTCTCGGATCCCAAGTCA 1488  
Qy 1426 ACCATTGTTGCCATCTTTCCTCCATCTCCCATCTTATATCTGCGCCCAACAGAGGTCAGTGG 1485  
Db 1489 ACCATTGTTGCCATCTTTCCTCCATCTTATATCTGCGCCCAACAGAGGTCAGTGG 1548  
Qy 1486 CACTGCAGGTCCTGGGATGATGCGGGTGCCCAATTTCTACATTTGCGGAGGGACCCGTGCA 1545  
Db 1549 CACTGCAGGTCCTGGGATGATGCGGGTGCCCAATTTCTACATTTGCGGAGGGACCCGTGCA 1608  
Qy 1546 GGAATGCCCATCTCTGAACCAAGAGGATCTGTATGAACCACTATGCGGGCAAGGTC 1605  
Db 1609 GGAATGCCCATCTCTGAACCAAGAGGATCTGTATGAACCACTATGCGGGCAAGGTC 1668  
Qy 1606 TTGAGCATGCCCTCGCCCTACCTCTGTGGAATCATTTCCATTCGCGAGTGGTGCCTAC 1665  
Db 1669 TTGAGCATGCCCTCGCCCTACCTCTGTGGAATCATTTCCATTCGCGAGTGGTGCCTAC 1728  
Qy 1666 AACAAAGCCAAAGAGGATGAGGATCTATGATCCAGCAGGACATGAGTTGACTTC 1725  
Db 1729 AACAAAGCCAAAGAGGATGAGGATCTATGATCCAGCAGGACATGAGTTGACTTC 1788  
Qy 1726 ATCTCAGGAACCTCGAATGAGGAGCTCGCCCGGAAGGAGAGAAATCCCCAGATGGCTTC 1785  
Db 1789 ATCTCAGGAACCTCGAATGAGGAGCTCGCCCGGAAGGAGAGAAATCCCCAGATGGCTTC 1848  
Qy 1786 ATGGCCCCAAAGCATGGAAGGTCCTGACAGATTTATACAGGTCCTGGGAGAGAACTAA 1845  
Db 1849 ATGGCCCCAAAGCATGGAAGGTCCTGACAGATTTATACAGGTCCTGGGAGAGAACTAA 1908

## RESULT 6

AF150754  
LOCUS 2377 bp mRNA linear PRI 25-JAN-2001  
DEFINITION Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthase 2b  
isoform mRNA, complete cds.  
ACCESSION AF150754  
VERSION AF150754.2 GI:12484558.

KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2377)

Venkatchalam, K.V., Fuda, H. and Strott, C.A.

3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform

Unpublished

2 (bases 1 to 2377)

Venkatchalam, K.V., Fuda, H. and Strott, C.A.

Direct Submission

Submitted (12-MAY-1999) Steroid Regulation, National Institute of

Health, 9000 Rockville Pike, Bethesda, MD 20892-4510, USA

3 (bases 1 to 2377)

Venkatchalam, K.V., Fuda, H. and Strott, C.A.

Direct Submission

Submitted (25-JAN-2001) Steroid Regulation, National Institute of

Health, 9000 Rockville Pike, Bethesda, MD 20892-4510, USA

Sequence update by submitter

On Jan 25, 2001 this sequence version replaced gi:6642928.

Location/Qualifiers

1. .2377

/organism="Homo sapiens"

/db\_xref="taxon:9606"

81..1940

/note="PAPS synthase 2b b; ATP sulfurylase/adenosine

5'-phosphosulfate kinase 2b; bifunctional enzyme that

regulates production of the universal sulfonate donor

molecule, 3'-phosphoadenosine 5'-phosphosulfate"

/codon\_start=1

/product="3'-phosphoadenosine 5'-phosphosulfate synthase 2b

isoform"

/protein\_id="AAF20366.2"

/db\_xref="GI:12484559"

## CDS

/translation="MSGIKKKQKTEKNOQKSTNVVQAHVSRNKRKQVVGTRGRRGCT  
VWLTGSGAGKTTISFALEYLVSHAIPTCYSLDGNVHGLNRLNGSPGDNENIR  
IAEVAKLIFADAGLVCTISFISPAFKDRENARKIHESAGLFFEFIFDPAENICSRD  
KGLYKRRAGEIKGTGIDSDYKPTPELVKLTNLSTSDCVHVOVLELQONIVPY  
TIIKDTHELEVPENKLDHVAEATLPSSLTKLDLQWVQLSGWATPLKGFNRKE  
YLOVMFDTLLDGMALPDGVNMSIPIVLPVSAEDTRLSGCKSFVLAHGRVAILR  
DABFEHRKEERSRVMTGTTCTKHPHKNHMGSDWLVDGLQVLEKIRWNGDLQYR  
LTFLELKCKCKEMNADAVAFQLRPNVHGHALLMODTRRLLERGTYPHVLPLHPLG  
GWTKDDVDVLDWRMKHAAVLEGLVDPKSTIVAFPSPLMYAGTPEVQWCHRSRLTA  
GANFYIVGRDPAGMPHPETKKDLYEPHTGKVLSPAGLTVEIIPRVAAYNKKAKA  
MDFYDPAHNFDFISGTRMRKLAREGENPDGFMAPKANVKLVLDIYRSLEKN"  
polyA\_site 2376  
BASE COUNT 636 a 556 c 584 g 601 t  
ORIGIN  
Query Match 98.3%; Score 1813.6; DB 9; Length 2377;  
Best Local Similarity 99.08; Pred. No. 0;  
Matches 1841; Conservative 0; Mismatches 4; Indels 15; Gaps 1;  
Qy 1 ATGTGCGGGATCAAGAAAGCAAAAGACCGGAGAACCCAGAGAACCCCAATGTAGTCTAT 60  
Db 81 ATGTGCGGGATCAAGAAAGCAAAAGACCGGAGAACCCAGAGAACCCCAATGTAGTCTAT 140  
Qy 61 CAGGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAACAAGGGTGGCTTC 120  
Db 141 CAGGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAACAAGGGTGGCTTC 200  
Qy 121 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGCTGCTGGAACAAACGATAAGTTT 180  
Db 201 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGCTGCTGGAACAAACGATAAGTTT 260  
Qy 181 GCCTCGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 240  
Db 261 GCCTCGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 320  
Qy 241 GTCCGTCATGCCCTTACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATC 300  
Db 321 GTCCGTCATGCCCTTACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATC 380  
Qy 301 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTCTGATGCTGCTGCTGCTGCATTACCAGC 360  
Db 381 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTCTGATGCTGCTGCTGCATTACCAGC 440  
Qy 361 TTTATTTCCTCATTCGCAAGGATCTGAGATGCGCGCAAAATACATGATCAGCAGG 420  
Db 441 TTTATTTCCTCATTCGCAAGGATCTGAGATGCGCGCAAAATACATGATCAGCAGG 500  
Qy 421 CTGCCATTCTTTCGAATATATTGTAGATGACCTCTAAATATTGTGAAGCAGAGACGTA 480  
Db 501 CTGCCATTCTTTCGAATATATTGTAGATGACCTCTAAATATTGTGAAGCAGAGACGTA 560  
Qy 481 AAAGGCTCTTATAAAAGGCGCAGAGTGGGAGATTAAAGGATTTACAGTATTGATTTCT 540  
Db 561 AAAGGCTCTTATAAAAGGCGCAGAGTGGGAGATTAAAGGATTTACAGTATTGATTTCT 620  
Qy 541 GATTATGAGAAACCTGAAACTCTCTGACCGTGTCTTAAACCAATTTGTCCACAGTACGT 600  
Db 621 GATTATGAGAAACCTGAAACTCTCTGACCGTGTCTTAAACCAATTTGTCCACAGTACGT 680  
Qy 601 GACTGTGTCACCCAGGTAGTGGAACTTCTGCAAGAGCAGAACATTTGTACCTTACTATA 660  
Db 681 GACTGTGTCACCCAGGTAGTGGAACTTCTGCAAGAGCAGAACATTTGTACCTTACTATA 740  
Qy 661 ATCAAGATATCCACGAACCTCTTTGTGCCGGAACAAACATTTGACCAAGTCCGAGCTGAG 720  
Db 741 ATCAAGATATCCACGAACCTCTTTGTGCCGGAACAAACATTTGACCAAGTCCGAGCTGAG 800  
Qy 721 GCTGAACCTCTCCTTCATTATCAATTAAGTCTGATCCAGTGGGTCAGGTTTGG 780  
Db 801 GCTGAACCTCTCCTTCATTATCAATTAAGTCTGATCCAGTGGGTCAGGTTTGG 860  
Qy 781 AGCGAAGGCTGGGCCACTCCCTCAAGAGTTTTCATCGGGAGAGGAGTACTTACAGGTT 840





[illegible]

Db 1754 TTTATGATCAGCAAGGACGAGGAGTTTGACTTCATCTCAGGAATCGCATGAGGAAGC 1813  
QY 1751 TCGCCCGGGAAGGAGAGAAATCCCCAGATGGCTTCATGGCCCCCAAGGATGGAAGTCC 1810  
Db 1814 TCGCCCGGGAAGGAGAGATCCCCAGATGGCTTCATGGCCCCCAAGGATGGAAGTGT 1873

QY 1811 TGACAGATTATTACAGTCCCTCGAGGAAGAACTA 1844  
Db 1874 TGACAGATTACTACAGTCTCTGGAGAGACCAA 1907

RESULT 9  
AF052453 1996 bp mRNA linear ROD 27-JUL-1998  
LOCUS Mus musculus ATP sulfurylase/APS kinase 2 mRNA, complete cds.  
DEFINITION AF052453  
VERSION AF052453.1 GI:3342265  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1996)  
AUTHORS Kurima, K., Warman, M.L., Krishnan, S., Domowicz, M., Krueger, R.C. Jr.,  
Deyrup, A., and Schwartz, N.B.  
TITLE A member of a family of sulfate-activating enzymes causes murine  
brachyorphism  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8681-8685 (1998)  
MEDLINE 98337975  
PUBMED 9671738  
REFERENCE 2 (bases 1 to 1996)  
AUTHORS Kurima, K.  
TITLE Direct Submission  
JOURNAL Submitted (04-MAR-1998) Kennedy Center, U. of Chicago, 5841 S.  
Maryland Ave. MC5058, Chicago, IL 60637, USA

## FEATURES

source  
1..1996  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
131..1996  
/EC\_number="2.7.7.4"  
/function="mediates two steps in sulfate activation:  
transfer of a sulfate group to ATP by ATP sulfurylase to  
yield adenosine 5'-phosphosulfate (APS), and the  
subsequent transfer of a phosphate group from APS to APS  
by APS kinase yielding ADP and  
3'-phosphoadenosine-5'-phosphosulfate (PAPS)"  
/note="bifunctional enzyme"  
/codon\_start=1  
/product="ATP sulfurylase/APS kinase 2"  
/protein\_id="AAC40191.1"  
/db\_xref="GI:3342265"  
/translation="MSANSKMKHNRDQKSTNNVYQAHVSRNKRQVVGTRGPRGC  
TWVLGSCAGATTTTSALEEVLSHAIPCYSLDGNVHRGLKNLGSFSGADREBNIR  
RAEVARARAGLVCITSFSPFKADRENARKIHESAGLPFEIFVDRPLNICESR  
VAGLVARADAGEIKGTGIDSYEKPTPECVKLNLSVSDCVQVQVBLLOEQNIVP  
HTYIKGLFELFVPEKVDQIRAEAEPLSPITKLDLOWQILSEGWATPLKFWREK  
EYLQTLFHTLDLVGVPRDVTNMSIPIVLVPSADDKARLEGCSKFAIMYEGRRVALL  
ODPEFVHRKEKRSRVWCTATAKPHIKMVMESGDLVYGGDLVLERIRWDDGLDQY  
RTPLELKCKCKMDNADVAFAQLRNPVNHGALLMODPRRLERGGYKHPVLLHPL  
GWTDDDDVPLWRKKQHAALVEERVLDPKSTIVAIFFSPMLYAGTEVQWHRCRMI  
AGANFVIGRDPAGMPHPTKKDLPEHGGKVLMSAPGLTSVEIIPFVRAAYNKIKK  
AMDYDPARHEFDFTSGTRMRKLREHGEDPPDGFMAPKRWLVLTIDYRSLEKTN"  
BASE COUNT 538 a 490 c 522 g 446 t

Query Match 76.2%; Score 1406.4; DB 10; Length 1996;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 1582; Conservative 0; Mismatches 251; Indels 15; Gaps 1;  
QY 12 CAAGAAGCAAAAGACGAGAACGACGAGAAATCCACCAATGTATCTATCAGGCCCAACCA 71

Db 145 CAAATGAACCATAAAGAGAGACCAGCAAAATCCACCAATGTGTCTTACCAGGCCCATCA 204  
QY 72 TGTGACGAGAAATAAGAGAGGCAAGTGGTTGGAACAAGGGGTGGGTTCGAGGATGTAC 131  
Db 205 TGTGACGAGAAACAAGAGAGCAAGTGGTTGGAACAGGGGAGGATTCGAGGATGTAC 264  
QY 132 CGTGTGGCTAACAGGTCTCTCTGTGTGCTGGAACCAACAGATAAGTTTTCCTCGGAGGA 191  
Db 265 CGTGTGGCTAACAGGTCTCTCTGTGTGCTGGAACCAACAGCATTAAGCTTTCTCTTGAAGA 324  
QY 192 GTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGTCATGG 251  
Db 325 GTACCTTGTATCTCAGGCCATCCCATGTTACTCTGTGATGGGACAAATGTCGTCATGG 384  
QY 252 CCTTAACAGAAATCTCGGATTCTCTCTGGGACAGAGAGAAATATCCCGCGGATTTC 311  
Db 385 CCTTAATAGAACTGGGATTCTCTGGGGGACCGAAGAGAGATATCCCGCGGATTCG 444  
QY 312 TGAGTGGCTTAAGCTGTTTGTCTGATGCTGTCTGGTCTGCATTACCAGCTTTATTCTCC 371  
Db 445 GGAGGTGGCCAGGCTCTTTGCCGACGCGGCTGGTTTGCATCACCAGCTTTATCTCTCC 504  
QY 372 ATTCGCAAGGATCGTGAGAAATCCCGCAAAATACATGAATCAGCAGGGCTGCCATCTT 431  
Db 505 TTTTGCAAAAGGATCGTGAGAAATCCCGCAAAATCCAGAAATCAGCAGGACTCCCGTCTT 564  
QY 432 TGAATATTGTAGATGCACCTCTAAATATTCTGAAAGCAGAGACGTAAAGGCCCTCTA 491  
Db 565 TGAGATCTTTGATAGATGCCCTTTAAATATCTGTGAAAGCCGAGAGCTAAAGGACTCTA 624  
QY 492 TAAAGGGCCAGAGCTGGGGAGATTAAAGGATTTACAGTATTGATTTCTGATTATGAGAA 551  
Db 625 CAAACGAGCCGAGAGAGAGATTAAGGGTTTACAGGCATCGATTCTCAGCTATGAGAA 684  
QY 552 ACCTGAAACTCTCGAGCTGTGCTTAAACCAATTTGTCCACAGTGTGAGTGTGTGCCA 611  
Db 685 ACCTGAAACTCCAGATGTGTGCTGAAGACCAACCTGTCTTCAAGTACGACCTGTGTGCA 744  
QY 612 CCAGGTAGTGGAACTTCTCAAGAGCAGAACATTTGACCTATACTATAATAAAGATAT 671  
Db 745 ACAGGTGGTGGAACTTTTGCAGGAGCAGAACATTTGACCCACACACCACCAAGGCAT 804  
QY 672 CCACGAACTTTTGTGCCGGAACAACTTGACCAGCTCCGAGCTCCGAGCTGAACTCT 731  
Db 805 CCACGAACTTTTGTGCCGGAACAACTTGACATCAAAATCCGAGCTCAGGACGAGACTCT 864  
QY 732 CCCTTCATTATCAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTAGCGAAGGCTG 791  
Db 865 CCATCACTACCAATTTACCAAGCTGGATCTGCAAGTGGGTGCGATTTCTGAGTGAAGGCTG 924  
QY 792 GGCCACTCCCTCAAAAGGTTTTCATCGGGGAGAAAGGAGTACTTACAGGTTATGCACTTTGA 851  
Db 925 GGCCACTCCCTCAAAAGGCTTTATCGGGGAGAAAGGAGTACTTGCACAACTCTACACTTCA 984  
QY 852 CACCCTGCTAGA-----TGATGGCTGTATCAACATGAGTATTTCCCTATTTGT 896  
Db 985 CACTCTACTGGAGCGGCTGTTCCCGCTGATGGAGTCATCAACATGATGATTTCCCTATTT 1044  
QY 897 ACTGCCCGCTCTCTGCAGAGGATAAGACACGCTGGAAGGTTGACAGAAAGTTTCTCTCGTGGC 956  
Db 1045 ATTGCCCGCTTTCGCGGATGACNAGGACGCGCTCGAAGGTTGCGAGCAAAATTTCCCTTGT 1104  
QY 957 ACATGGTGACGAGGAGTGTAGCTATCTTACGAGACGCTGAATTTCTATTAACACAGAAAGA 1016  
Db 1105 GTACGAAGTTCGAGGGTCCGCTCTATTACAGGACCCCTGAATTTCTATGAGCATAGGAAGA 1164  
QY 1017 GGAACGCTGTTCCGCTGTTTGGGGACACATGTACAAACACCCCCCATATCAAAATGGT 1076  
Db 1165 GGAGCGCTTCTTCGTTGTGGGGAAACAGCCATTCGAAAGCAGCCCATATCAAAATGGT 1224  
QY 1077 GATGAAAGTGGGAGCTGGCTGTTGTTGAGACCTTCAGGTGCTGAGAGAAATAAGATG 1136

```

Db 1225 GATGAAAGTGGGACTGGCTTGTGGTGGAGACCTACAGGTCTAGAGAGAAATPAGGTG 1284
QY 1137 GAATGATGGGTGGACCAATACCGCTCTGACACCTCTGGAGCTCAACACAGAAATGTAAGA 1196
Db 1285 GGACATGGGTGGACCAATACCGCTCTACGCTCTGGAGCTCAACACAGAAATGTAAGA 1344
QY 1197 AATGAATGCTGATCGGTGTGGTTCATTCAGTTCGCAATCTGCTCCACAAATGGCCATGC 1256
Db 1345 CATGAATGCTGATCGGTGTGGTTCATTCAGTTCGCAATCTGCTCCACAAATGGTCAATGC 1404
QY 1257 CTTCTTTGATGAGGACACCTCGCGAGGCTCTAGAGAGGGCTACAAAGCACCCGCTCT 1316
Db 1405 CTTCTGATGAGGACACCCGCGAGGCTCTGAGAGGGTTACAGACCCAGCTCT 1464
QY 1317 CCTACTACACCTCTGGGCGGTGGACCAAGATGACGATGTCCTCTAGACTGGCGGAT 1376
Db 1465 CTTGCTCACCTCTTGGGGGTGGACCAAGGACGATGACGCTACCTCTGGAATGGAGAT 1524
QY 1377 GAACAGCACCGGCTGCTCGAGGAGGGTCTGTGATCAACCTCAACCTATGTTGC 1436
Db 1525 GAACAGCATGACGCTGACTGGAGGAAAGGTTCTGTGATCCCAAGTCAACTATTGTTGC 1584
QY 1437 CATCTTTCCGCTCTCCATGTTATGCTGGCCCCACAGAGGTCCAGTGGCCTGCGAGTCT 1496
Db 1585 CATCTTTCCATCTCTATGTTATGCTGTGCTCCACAGAGTCCAGTGGCATGTCAGATG 1644
QY 1497 CCGGATGATGCGGTGGCAATTTCTACATTTGTTGGGAGGACCTCGAGGAATGCCCA 1556
Db 1645 CCGGATGATGCGAGGACCAATTTCTACATTTGTTGGTAGGATCCCGCAGGAATGCCCA 1704
QY 1557 TCCTGAACCAAGAGGATCTGTATGATCAACCTCACTATGGGGCAAGGTCTTGAGCATGGC 1616
Db 1705 TCCTGAGACAAAGAAAGACCTATATGAACCCACCGGGGCAAGGTCTTGATGATGGC 1764
QY 1617 CCTGGCTCACCTCTGTGGAATCATTTCCAGTTCGAGTGGCTGCTTACACAAAGCCAA 1676
Db 1765 CCTGGCTCTACCTCTGTGGAATTAATTCGTTCCGAGTGGCTGCTTACATAAATTA 1824
QY 1677 AAAAGCCATGAGCTCTATGATCCAGCAAGGACCAATGAGTTGACTTCTATCTCAGGAAC 1736
Db 1825 AAAGCCATGAGCTTTATGATCCAGCAAGGACCAAGGAGTTGACTTCTATCTCAGGAAC 1884
QY 1737 TCGATGAGGAGCTGCGCCGGGAGGAGAGAAATCCCCAGATGCTTCATGGCCCCCAA 1796
Db 1885 TCGATGAGGAGCTGCGCCGGGAGGAGAGAAATCCCCAGATGCTTCATGGCCCCCAA 1944
QY 1797 AGCATGGAAGTCTCTGACAGATTATTACAGGTCCCTGGAGAGAACTA 1844
Db 1945 AGCTGGAGAGTGTGACAGATTACTACAGTCTCTGGAGAGACCAA 1992

RESULT 10
MMU34883
LOCUS MMU34883 2479 bp mRNA linear ROD 08-DEC-1995
DEFINITION Mus musculus ATP sulfurylase/APS kinase mRNA, complete cds.
ACCESSION U34883
VERSION U34883.1 GI:1109675
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2479)
Li, H., Deyrup, A., Mensch, J. R. Jr., Domowicz, M., Konstantinidis, A. K.
and Schwartz, N. B.
The isolation and characterization of cDNA encoding the mouse
bifunctional ATP sulfurylase-adenosine 5'-phosphosulfate kinase
J. Biol. Chem. 270 (49), 29459 (1995)
96094345
PUBMED 7493984
REFERENCE 2 (bases 1 to 2479)
AUTHORS Mensch, J. R.
TITLE Direct Submission

```

Submitted (25-AUG-1995) James R. Mensch, Pediatrics, University of Chicago, 5841 S. Maryland Ave. MC5058, Chicago, IL 60637, USA

Location/Qualifiers

1. 2479

/organism="Mus musculus"

/db\_xref="taxon:10090"

34. 1908

/note="Bifunctional enzyme; mediates two steps in sulfate activation: transfer of a sulfate group to ATP by ATP sulfurylase (ATP sulfate adenylyltransferase, EC 2.7.7.4) to yield adenosine 5'-phosphosulfate (APS), and the subsequent transfer of a phosphate group from APS to APS by APS kinase (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase, EC 2.7.1.25) yielding ADP and 3'-phosphoadenosine-5'-phosphosulfate (PAPS)"

/codon\_start=1

/product="ATP sulfurylase/APS kinase"

/protein\_id="AAC52328.1"

/db\_xref="GI:1109676"

/translation="MEIPGSLCKVKLSNNAQNMQRATNVTYQAHVSRNKRGOVY  
TRGGRFCTVMTLGLSGAGKTTVSMALVEEYLVCHGIPCYTLDGNIHQGLNKLIFS  
PEDRENVRIAEVAKLFADAGLVCTISFYPTQDRNARQIHGASLSLPFEVFWDA  
PLHCEQRDVGLYKARAGEIKGFTGIDSEYKPEAPELVLTCDNDVQVQVVE  
LLOERDIYPVDASVEYKELYVPENKHLAKTDAEALPALKINKVDQWQVLAESWAT  
PLNGFMREYLDQCLHFDLLDGGVNLSPVILVATHEKREDKLDGCTAFALVIGRR  
VALLRNPEFEHREKRCAROMGTTCKNHPIKMYLEQDMLIGDQLQVLDRIWNG  
LDQVRLTPTLTKQKFKDMNDAVAFQLRNPVNHGHALMODTHKQLLERYRRPVL  
LHPILGGTDDDDPLMWRMHPAALEBILDPETTVVAIFSPMVMYAGTPVRAHYN  
ARMVAGANFYIVGRDPAGMHPETGKDLIEPTHGAKVLTMAPGLTLEIVPVRVACR  
YKKRMDYDSEHEDHDFEISITRMRKLAREGQKPEPGEFMAPKAWTVLVEYKSLKA  
"

polyA\_site 2479

BASE COUNT 613 a 583 c 667 g 616 t

ORIGIN

Query Match 54.5%; Score 1005; DB 10; Length 2479;

Best Local Similarity 71.7%; Pred. No. 4.5e-276;

Matches 1317; Conservative 0; Mismatches 520; Indels 0; Gaps 0;

QY 9 GATCAAGAGCAAAAGACGGAGAACCCAGCAGCAAAATCCACCAATGATGTCTATCAGGCCCA 68

Db 72 GAGCAATTAACGGCGAGAACTGGGGATGCGAGAGACCAACCACTACCTATCAAGCCCA 131

QY 69 CCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGACAAAGGGGTGGTTCGCGGATG 128

Db 132 CCATGTGAGCAGGAATAAGAGAGGCGAGTGTGGGGACAGAGTGGCTCCGTTGGTTG 191

QY 129 TACCGTGTGGCTAACAGGTCTCTCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 188

Db 192 CACAGTTTGGCTGACAGGATTTGTCGGGAGCAGGAGAGACGCGGTGAGCATGGCTCTGA 251

QY 189 GGAGTACCTGTCTCCCATGCTCCCTGTGTTACTTCTGATGGGAGCAATGTCGCTCA 248

Db 252 GGAGTACCTGTGTCGCCAGGATTCGCTGCTACACTTTGGATGTTGACACATCGCCA 311

QY 249 TGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAGAAATATCCGCGGAT 308

Db 312 AGGACTCAATAAGAACCTCGGCTTCAGTCTCGAGCAGAGAGAGAGAGAGGATTCGCGCAT 371

QY 309 TGCTGAGTGGCTAGCTGTTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368

Db 372 AGCTGAGTGGCGAAGCTGTTTGGCAGATGCTGGCTTAGTGTATCACCAGCTTTATATC 431

QY 369 TCCATTCGAAGGATCTGAGAAATCCCGCAAAATACATGATCAATCAGAGGCTGCCATT 428

Db 432 GCTTTACACAGGATCGCAACACAGCAAGGAGATTCATGAGGTGCAAGCTTCCCTT 491

QY 429 CTTTGAAATATTTGTAGATGCACTTCAATATTTGTGAAAGCAGAGAGAGGCTTAAAGGCT 488

Db 492 CTTTGAGATTTTGTGATGCTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551

QY 489 CTATAAAGGCCCGAGCTGGGAGATTTAAAGGATTTACAGGATTTGATTTCTGATATGA 548



|    |      |   |      |
|----|------|---|------|
| DB | 552  | CTACAAGAGCGCGGCGAGATAAAGGCTTCACGTGGCATGATTCTCAGTATGA    | 611  |
| QY | 549  | GAACCTTGAACCTCCTGAGGCTGTCTTAAACCAATTTGTCCACAGTGAAGTGTGT | 608  |
| DB | 612  | GAACCCGAGGCGCCGAGCTGTGCTGAAACGGATTCTCTGACGCTCAACGACTCGT | 671  |
| QY | 609  | CCACGAGTAGTGGAACTTCTGCNAGAGCAGACATTGTACCTTATCTATATCAAGA | 668  |
| DB | 672  | CCAGCAGGTTGTGGAGCTTCTCAGGAACGGGACATCGTCCCTGTGGATGCTTCC  | 731  |
| QY | 669  | TATCCAGCAACTCTTTGTGCGGAAAAACAACCTTGACACGTCGAGCTGAGGCTGA | 728  |
| DB | 732  | AGTGAAGAGCTATATCTGCCAGAGATAAATTCATTGCGCCAAACTCATGCAGA   | 791  |
| QY | 729  | TCTCCCTTCATATCAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTGAGCAG  | 788  |
| DB | 792  | CTTACCAGCCCTGAAATCAATAAGTGGATATGACGTGGGTGAGGTTTGGCAGA   | 851  |
| QY | 789  | CTGGGCCACTCCCTCAAAGGTTTCATGCGGAGAGGAGTACTTACAGGTTATGCA  | 848  |
| DB | 852  | TTGGGCCACTCTCTGAACGCTTCATGAGAGAGGAGTACTTGCAGTGCCTCCAT   | 911  |
| QY | 849  | TGACACCCCTGCTAGATGATGGCTGTATCAACATGAGCATCCCCATTTGATGCC  | 908  |
| DB | 912  | CGATTGTCTTGTGATGGAGCGCTATCAACTTATCGTGGCTATAGTTCTGACAG   | 971  |
| QY | 909  | TGCAGAGATAAGACACGCTGGAAGGTCACAGCTTGTCTCGACCATGGTGGACG   | 968  |
| DB | 972  | GCACGAGATAAGAGAGAGCTTGGACGCTGCACCGGCTTCGCTGTGTATGAGG    | 1031 |
| QY | 969  | GAGGTTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGAACG   | 1028 |
| DB | 1032 | CCGCGTGGCCATCTTCCGAATCTGAATTTTGTGACACGGAAAGAGGCGTGC     | 1091 |
| QY | 1029 | CCGTGTTGGGGACAACATGTACAACACCCCCCATATCAAAATGCTGATGNAAG   | 1088 |
| DB | 1092 | CAGACAGTGGGAAACAATGTCAAGAACCCACCCCTACATCAAGATGGTTCTG    | 1151 |
| QY | 1089 | GGACTGGCTGTGGTGGAGACCTTCAAGTGTCTGGAGAAAATAGATGGAATGAT   | 1148 |
| DB | 1152 | GGATTGCTGATTGGAGSAGATCTTCAAGTCTTGGACCGGATTACTTGAATGAT   | 1211 |
| QY | 1149 | GGACCAATACGCTTGACACTCTGGAGCTCAAAACAGAAATGTAAAGAAATGA    | 1208 |
| DB | 1212 | TGATCAGTACCGCTTACCCGACGGAGCTCAAGCAGAAGTTTAAAGATATGAC    | 1271 |
| QY | 1209 | TGCGGCTTTTGCAATTCAGTTGGCAATCTGTGCCAATGGCATGCCCTGTGTATG  | 1268 |
| DB | 1272 | TGCTGCTTTTGCAATTCAGTTGGCCACACCAAGTGCACACGGGACGCTCTGT    | 1331 |
| QY | 1269 | GGACACCTCGCAGGCTCTTAGAGAGGGCTACAAGCACCCGCTCTCTACTACAC   | 1328 |
| DB | 1332 | GGATACCCAAAGCAGCTTCTGGAGAGGGCTACCGCGCCCTGTCTCTCTCCAT    | 1391 |
| QY | 1329 | TCTGGGCGGCTGGACCAAGATGACGATGTCCCTCTAGACTGGCGGATGAAGCA   | 1388 |
| DB | 1392 | TCTTGTGGCTGGACGAAGATGACGATGTCCCTCTGATGTGGCGTATGAAGCA    | 1451 |
| QY | 1389 | GGCTGTCTCGAGAGGGGTCTTGGATCCCAAGTCAACCATTTGTTGCCATCTTT   | 1448 |
| DB | 1452 | TGCAGTCTGGAGAGGSCATCTGGATCCTGAACGACAGTGGTGCCATTTCCG     | 1511 |
| QY | 1449 | TCCCATGTTATATGCTGGCCCCACAGAGGTCCAGTGGCATGCGAGTCCCGAT    | 1508 |
| DB | 1512 | TCCTATGATGATGCTGGGCCAACCGAGGTCCAGTGGCACTGCGAGGCGGAT     | 1571 |
| QY | 1509 | GGGTGCCAATTTCTACATTTGTGGGAGGGACCTGCAGGAATGCCCATCTCT     | 1568 |
| DB | 1572 | CGNAGCCAAATTTTATATTGTTGGACGAGACCCCTGCTGGCATGCCTCAT      | 1631 |
| QY | 1569 | GAAGGATCTGTATGAACCCACTCATGGGGCAAGGTTTGAAGCATGCCCTGC     | 1628 |

|                       |  |  |                    |
|-----------------------|--|--|--------------------|
| Db                    | 1632   | GAAGGACCTCTATGAGCGCAACACATAGTGTGCCAAAGTCTGACGATGCCCCAGCCCTGAT  | 1639               |
| QY                    | 1629   | CTCTGTGGAAATCATTTCCATTCAGAGTGGTCCCTACACAAAGCCAAAAAGCCATGGA     | 1688               |
| Db                    | 1692   | TACCTTTGGAAATCGTTCCCTTCCAGTTCAGCTTTACACAAAGAAAAAGACGGATGGA     | 1751               |
| QY                    | 1689   | CTTCTATGATCCACGACGACATGAGTTTGACTTTCATCTCAGGAAGTCTCAGTGAAGAA    | 1748               |
| Db                    | 1752   | CTACTATGACTCTCAGGACCAACGAGACTTCGAGTTTATTTCAGGAACGAGGATGCGCAA   | 1811               |
| QY                    | 1749   | GCTGCCCGGGAGAGAGAAATCCCCAGATGGCTTCATGGCCCCCAAGCATGGAAGGT       | 1808               |
| Db                    | 1812   | GCTGSCACGGAAGCCAGAAACCTCCAGAGGGCTTCATGGCCCCCAAGGCTTGGAGTGT     | 1871               |
| QY                    | 1809   | CCTGACAGATTTATACAGGTCCCTGGGAGAGAACTAA                          | 1845               |
| Db                    | 1872   | GCTGTTAGAGTACTACAAGTCTTTAGAGAAAGCCTAA                          | 1908               |
| RESULT                | 11   |  |                    |
| E27970                |  |  |                    |
| LOCUS                 | E27970   | 1875 bp DNA linear   | PAT 18-JUN-2001    |
| DEFINITION            | APS kinase/ATP sulfurylase gene derived from human being.  |  |                    |
| ACCESSION             | E27970   |  |                    |
| VERSION               | E27970.1   | GI:13026556  |                    |
| KEYWORDS              | JP 1999187883-A/1.   |  |                    |
| SOURCE                | Homo sapiens.  |  |                    |
| ORGANISM              | Homo sapiens.  |  |                    |
| REFERENCE             | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  |  |                    |
| AUTHORS               | 1 (bases 1 to 1875)  |  |                    |
| TITLE                 | Katsuhiko,T., Hiroshi,N. and Masahito,S.   |  |                    |
| JOURNAL               | APS kinase/ATP sulfurylase gene derived from human being   |  |                    |
| COMMENT               | Patent: JP 1999187883-A 1 13-JUL-1999;<br>JAPAN HEALTH SCIENCES FOUNDATION,UNITIKA LTD<br>OS Homo sapiens (human)<br>PN JP 1999187883-A/1<br>PD 13-JUL-1999<br>PF 26-DEC-1997 JP 1997360387<br>PR<br>PI KATSUHIKO TAKAYANAGI, HIROSHI NAKAJIMA, MASAHITO SUIKO PC<br>C12N15/09, C12N1/21, C12N9/12, C12P17/18, (C12N15/09, C12R1:91), PC<br>(C12N1/21, C12R1:19), (C12N9/12, C12R1:19), C12N15/00, (C12N15/00, PC<br>C12R1:91)<br>CC Strandedness: Double;<br>CC Topology: Linear;<br>FH Key<br>FT source<br>FT Location/Qualifiers<br>1..1875<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>536 a 379 c 499 g 471 t |  |                    |
| FEATURES              | source   |  |                    |
| BASE COUNT            | 536 a 379 c 499 g 471 t  |  |                    |
| ORIGIN                |  |  |                    |
| Query Match           | 53.7%;   | Score 990;   | DB 6; Length 1875; |
| Best Local Similarity | 71.3%;   | Pred. No. 8.6e-272;  |                    |
| Matches 1305;         | Conservative 0;  | Mismatches 525;  | Indels 0; Gaps 0;  |
| QY                    | 9  | GATCAAGAAGCAAAAGACGGAGAACACAGAGAAATCCACCAATGTACTCTATCAGGCCCA   | 68                 |
| Db                    | 39   | GAGCAATACGCCGAGAACTGGGAATGCAGAGAGCAACCAANTGTACCTTACCAGGCCCA    | 98                 |
| QY                    | 69   | CCATGTCGACGAGGAATAAGAGAGGGCAAGTGGTTGGAAACAGAGGGTGGGTTCCGAGGATG | 128                |
| Db                    | 99   | TCATGTCGACGAGGAACAAAGAGAGGTTCAGGTGGTGGGACCAAGAGTGGCTTTGCTGGTTG | 158                |
| QY                    | 129  | TACCGTGTGGCTACAGGTCTCTCTGCTGCTGGAAAAACACAGATAGTTTTCGCCCTGGA    | 188                |
| Db                    | 159  | CACAGTTTGGCTAACAGGCTTGTCTGGAGCGGGGAAAGACTACTGTGAGCATGGCCTTGGGA | 218                |
| QY                    | 189  | GGAGTACTTGTCTTCCCATGCCATCCCTTGTACTCTCCCTGGATGGGAGCAATGTCCGTCA  | 248                |



|            |  |  |        |
|------------|--|--|--------|
| Db         | 1399   | GGATACCCATAAGCAACTTCTAGAGAGGGGCTACCGGGCCCTGTCTCTCTCTCCACCC     | 1358   |
| Qy         | 1329   | TCGGGGGGCTGGACCAAGGATGACGATGTGCCCTCTAGACTGGCGGATGAAGCAGACGC    | 1388   |
| Db         | 1359   | TCGGGGGGCTGGACCAAGGATGACGATGTCTCTGATGGCGTATGAAGCAGCATGC        | 1418   |
| Qy         | 1389   | GGCTGTGCTCGAGGAAGGGTCTCTGGATCCCAAGTCAACCATTTGTTGCCATCTTTCCGTC  | 1448   |
| Db         | 1419   | TGCAGTCTTGGAGGAAGGAGTTCTGAATCTCTGAGACGACAGTGGTGCCATCTTCCCATC   | 1478   |
| Qy         | 1449   | TCCCATGTTATAGCTGGCCCCCAGAGGTCCTGATGAGTGCAGCTGCAGGTCCCGGATGATGC | 1508   |
| Db         | 1479   | TCCCATGATGATGCTGGACCAACTGAGGTCCAGTGGCATTGCAGAGCAGCGATGGTTGC    | 1538   |
| Qy         | 1509   | GGGTGCCAAATTTCTACATTTGTGGGAGGGACCTTCAGCAAGTATGCCCATCTCTGAACCAA | 1568   |
| Db         | 1539   | AGGAGCCAACTTTTACATTTGTCAGCAGACCCCTGCTGGCATGCTCATCCAGAAACAGG    | 1598   |
| Qy         | 1569   | GAAGGATCTGTATGAACCCACTCATGCGGGCAAGGCTTGTGAGCATGCCCTTGGCCCTCAC  | 1628   |
| Db         | 1599   | GAAGGATCTTTATGAGCCAAGTCATGCTGCCAAAGTCTGACGATGGCCCTGGTTTAAT     | 1658   |
| Qy         | 1629   | CTCTGTGGAATCATTTCCATTCAGTCCAGTGGCTGCTTACAAACAAAGCCAAAAGCCATGGA | 1688   |
| Db         | 1659   | CACTTTGGAAATAGTTCCCTTTCAGTTGCGACTTACAAAGAAAAGACGATATGGA        | 1718   |
| Qy         | 1689   | CTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTTCATCTCAGCAACTCGAATGAGAA  | 1748   |
| Db         | 1719   | CTACTATGACTCTGAACCACTGAAGACTTTGANTTTATTTTCAGAACAGGAATGCCAA     | 1778   |
| Qy         | 1749   | GCTCGCCCCGGAGAGAGGAATCCCCAGATGGCTTCATGCCCCCAAGCATGGAAGGT       | 1808   |
| Db         | 1779   | ACTTGCTCGAGAAGGCCAGAAACACCATGAAGGTTTCATGGCTCCCAAGGCTTGGACCGT   | 1838   |
| Qy         | 1809   | CCTGACAGATTTATACAGGTCCTCTGGAGAA                                | 1838   |
| Db         | 1839   | GCTGACAGAATACTACAAATCCTTTGGAGAA                                | 1868   |
| RESULT 12  |  |  |        |
| AF033026   |  |  |        |
| LOCUS      | AF033026   | 2265 bp  | linear |
| DEFINITION | Homo sapiens bifunctional ATP sulfurylase/adenosine                    |  |        |
| ACCESSION  | AF033026   | 5'-phosphosulfate kinase mRNA, complete cds.                   |        |
| VERSION    | AF033026.1   | GI:3378100   |        |
| KEYWORDS   |  |  |        |
| SOURCE     | Homo sapiens.  |  |        |
| ORGANISM   | Homo sapiens   |  |        |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;      |  |        |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.             |  |        |
| TITLE      | 1 (bases 1 to 2265)  |  |        |
| JOURNAL    | Yanagisawa, K., Sakakibara, Y., Suiko, M., Takami, Y., Nakayama, T.,   |  |        |
| MEDLINE    | Nakajima, H., Takayanagi, K., Natori, Y. and Liu, M.-C.                |  |        |
| PUBMED     | cDNA cloning, expression, and characterization of the human            |  |        |
| AUTHORS    | bifunctional ATP sulfurylase/adenosine 5'-phosphosulfate kinase        |  |        |
| TITLE      | enzyme   |  |        |
| JOURNAL    | Biosci. Biotechnol. Biochem. 62 (5), 1037-1040 (1998)                  |  |        |
| MEDLINE    | 98312048   |  |        |
| PUBMED     | 9648242  |  |        |
| AUTHORS    | 2 (bases 1 to 2265)  |  |        |
| TITLE      | Yanagisawa, K., Sakakibara, Y., Suiko, M., Nakayama, T., Nakajima, H., |  |        |
| JOURNAL    | Natori, Y. and Liu, M.-C.  |  |        |
| FEATURES   | Direct Submission  |  |        |
| SOURCE     | Submitted (05-NOV-1997) Biochemistry, University of Texas Health       |  |        |
|            | Center at Tyler, P.O. Box 2003, Tyler, TX 75710, USA                   |  |        |
|            | Location/Qualifiers  |  |        |
|            | 1. .2265   |  |        |
|            | /organism="Homo sapiens"   |  |        |
|            | /db_xref="taxon:9606"  |  |        |
|            | /tissue_type="brain"   |  |        |
|            | /dev_stage="infant"  |  |        |
|            | 27. .1901  |  |        |
| CDS        |  |  |        |

27. .1901  
/dev -slay

```
/codon_start=1
/product="bifunctional ATP sulfurylase/adenosine
5'-phosphosulfate kinase"
/protein_id="AAC28429.1"
/db_xref="GI:3378101"
/translation="MELPSGLCKKYLNNQWGMORATNVTYQAHVSRNKGQV
GTRGFGCTVWLTGSGAKTIVSMALEYLVCHGIPCYTLDGDNIRQGLNKLGF
PEDREENVRIAGYAKLIFADAGLVCIITFSIPYODRNARQIHEGASLPFFVYD
PLHCEORDVKGILYKARAGEIKGFTGIDSEYKPEAPELVKLTDCDVNDVQVVE
LLOERDITVPDASVEVELYVPENKHLAKTDAETLPALKINKVMOWOVLAEGNAT
PLANGFEREYVLOCLHEDCLLDGGINLSVPIVLTAHEDKERLDSCTAFALMYEGR
VAILRNPFTEHKEERCARQWGTCKNHPYIKMVEQDGLGGDLQVLDVRWINDG
LDQRLNPFTEHKEERCARQWGTCKNHPYIKMVEQDGLGGDLQVLDVRWINDG
LHPLGWKTDKDDVPLMWRKQHAHAEVLEENETTVVAFPPSPMYAGPTEVQWHR
ARVAGANFYIVRGDPAGMPHETKDLYPEHSHGKVLITNAPGLITLIEIVPFRVAYN
KKKRMDDYDSEHHEDEFTISGTRMRKLAREGQKPPGEPFNAKWTVLTYEYKLEKA
"
BASE COUNT      628 a  458 c  559 g  620 t
ORIGIN
Query Match      53.7%; Score 990; DB 9; Length 2265;
Best Local Similarity 71.3%; Pred. No. 8.8e-272;
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;
QY 9 GATCAAGAACAAAGAGGAGAACAGACAGAGAAATCCACCATGTAGTCTATCAGGCCCA 68
DB 65 GAGCAATTAACGCCGAGCAACTGGGGAATGCAGAGAGCAACCAATGTCACTTACCAAGCCCA 124
QY 69 CCATGTGAGCAGGAAATGAAGAGAGGCAAGTGGTGGAAACAAGGGGTTCGAGGATG 128
DB 125 TCATGTGACGAGGACAGAGAGAGGTGAGGTGGTGGGACACAGAGGTGGCTTTGCGTTG 184
QY 129 TACCGTGTGCTAACAGGTCTCTGTGTGCTGCTGAAAAACAACGATAAGTTTTGCCCTGGA 188
DB 185 CACAGTTTGGCTTAACAGGCTTGTCTGAGCGGGAAGAGACTACTGTGAGCATGCCCTTGA 244
QY 189 GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGAGCAATGTCCGTCA 248
DB 245 GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGAGCAATGTCCGTCA 304
QY 249 TGGCCTTAAACAGAAATCTCGGATTTCTCCTCGGGACAGAGAGAAATATCCGCGGAT 308
DB 305 AGGTCTCAATAAATACTTGGCTTAGTCTGAGACAGAGAGAGATGTTTCGACGCAT 364
QY 309 TGCTGAGGTGGCTAAGCTGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
DB 365 CGCAGAGAGTTGCTAAACTGTTTGCAGATGCTGGCTTAGTGTGCATCAAGAGTTTCATATC 424
QY 369 TCCATTGCAAGAGATGCTGAGATGCCCAAAATACATGAATCAGCAGGGCTGCCATT 428
DB 425 ACCTTACACTCAGATCGCAACATGCAAGGCAATTCATGAGGTCAGAGTTTACCGTT 484
QY 429 CTTTGAATATTGTAGATCACTCTAAATATTGTGAAAGCAGAGACGTAAGAGCCT 488
DB 485 TTTTGAAGTATTGTTGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
QY 489 CTATAAAGGCCAGAGCTGGGAGATTAAGAGATTTACAGGATTTGATTTCTGATTAAGA 548
DB 545 CTACAAAAGGCCAGGAGAGAAATTAAGGTTTCACTGGGATCGATTTCTGAATATGA 604
QY 549 GAAACCTGAAACTCTTACGAGCTGTGCTTAAACCAATTTGCTCCACAGTGAAGTCACTGTGT 608
DB 605 AAGCCAGAGGCCCTTGAGTTGGTGTGTAAGAACAGACTCTCTGTGATGAATGACTGTGT 664
QY 609 CCACAGGTAGTGAAGTCTTCAAGAGCAGAGACATTTGTACCCCTATATAATCAAGA 668
DB 665 CCAGCAAGTTGTGAAGTCTTACAGGACGGGATTTGTACCTGTGATGCTATTAAGA 724
QY 669 TATCCAGGAACTCTTGTGCGGGAACAACATTTGACAGCTCCGAGCTCGAGGCTGAAC 728
DB 725 AGTAAAGAACTATATGTGCGCAGAAATAAATCTTCACTTTTGGCAAAAACAGATCGGAAAC 784
QY 729 TCTCCCTTCATTATCAATTACTAAGCTGGATCTCCAGTGGGTCCAGTTTGTAGCGAAG 788
```

```
DB 785 ATTACCAGCACTGAAAAATTAATAAGTGGATATCAGTGGGTGCAGGTTTGGCAGAAG 844
QY 789 CTGGCCACCTCCCTCAAGAGTTTTCATGCGGAGAGAGTACTTACAGGTTATGCACTT 848
DB 845 TTGGGCAACCCCATTTGAATGGCTTTATGAGAGAGAGGAGTACTTTCAGTGCCTTCATTT 904
QY 849 TGACACCCCTCTAGATGATGCGGTGATCAACATCAGCATGCCATCTTACTGCCCGCTC 908
DB 905 TGATTGCTCTTGATGAGGTGTCATTAACTTGTCACTACCTATATCTTCTGACTGCGAC 964
QY 909 TGCAGAGGATAAGACACGGCTGGAAGGGTGCAGCAAGTTTGTCTGCGACATGGTGGAGC 968
DB 965 TCATGAAGATAAAGAGAGGCTGGACGGCTGTACAGCATTTTCTCTGATGTATGAGGCGC 1024
QY 969 GAGGTGAGTATCTTACGACACGCTGAATTTCTATGAACACAGAAACAGAACGCTCTTC 1028
DB 1025 CCGTGTGGCCATCTTCGCAATCCAGAGTTTTTGGACACAGGAAAGAGAGCGCTGTC 1084
QY 1029 CCGTGTGTGGGACAAACATGTACAAACACCCCATATCAAAATGTGTATGGAAGTGG 1088
DB 1085 CAGACAGTGGGGAACGACATGCAAGAACACCACCTATATTAAAGATGCTGATGGAACAAG 1144
QY 1089 GGACTGGCTGGTGGTGAGACCTTCAAGTCTTGGATCGAGTTTATTGSAATGATGGCT 1148
DB 1145 AGATTGGCTGATTGGAGGAGATCTCAAGTCTTGGATCGAGTTTATTGSAATGATGGCT 1204
QY 1149 GGACCAATACCGTCTGACACCTCTGGAGCTCAACACAGAAATGTAAAGAAATGAATGCTGA 1208
DB 1205 TGATCAGTATCGTCTTACTCTACTGAGCTTAAGCAGAAATTTAAAGATATGATGCTGA 1264
QY 1209 TGGCGGTGTTGCAATCCAGTTGGCCAATCTGTCTCCACAATGGCCATGCCCTGTGTGATGCA 1268
DB 1265 TGTGCTTTTGGCATTTCACTACGCAACCCAGTGCACAATGGACATGCCCTGTGTAATGCA 1324
QY 1269 GGACACTGCCGAGGCTCTAGAGAGGGCTACAAGCACCCGGTCTCTCTACTACTACACC 1328
DB 1325 GGATACCCATAAGCAACTTCTAGAGAGGGGCTACCGGCGCCCTGTCTCTCTCTCCACCC 1384
QY 1329 TCTGGCGGCTGACCAAGGATGACGATGTGCTCTAGACTGGCGGATGAAGCAGACGC 1388
DB 1385 TCTGGGTGGCTGGCAAAAGGATGACGATGTCTCTTGATGTGGCGTATGAAGCAGCATGC 1444
QY 1389 GGCTGTGCTCGAGAGAGGGTCTCTGGATCCCAAGTCAACCATTTGTGCGCATCTTTCGGTC 1448
DB 1445 TGCAGTGTGGAGGAGAGGTTCTGAATCTCTGAGACGACAGTGTGTGCGCATCTTCCCATC 1504
QY 1449 TCCCATGTTATATGCTGGCCCCCAGAGGTCAGTGGCAGTCCGAGGTCCTCGGATGATTC 1508
DB 1505 TCCCATGATGTATGCTGGACCAACTGAGGTCCAGTGGCATTTGCAGACGCGATGTTGTC 1564
QY 1509 GGTGTCCTCAATTTCTACATTTGGGAGGGACCCCTGCAGGAATCCCCATCTCTGAAACCA 1568
DB 1565 AGGAGCCAACTTTTACATTTGAGCAGACCCCTGCTGGCATGCTCTCATCCAGAACAGG 1624
QY 1569 GAAGGATCTGTATGAACCCCACTCATGGGCGCAAGTCTTTCAGCATGCCCCCTGGCCTCAC 1628
DB 1625 GAAGGATCTTTATGAGCCAAAGTCTGTTGCCAAGTGTGACGATGCCCCCTGTTTAAAT 1684
QY 1629 CTCTGTGGAATCATTTCCATTTCCGAGTGGCTGCTTACAACAAAGCCAAAAAGCCATGGA 1688
DB 1685 CACTTTTGAATAAGTTTCCCTTTTCGAGTTGCGAGTTTACACAGAAAGAAAGAGCGTATGA 1744
QY 1689 CTTCTATGATCCAGCAGGCAATGAGTTTGTACTTCTATCTCAGGAACCTCGAATGAGGAA 1748
DB 1745 CTACTATGACTCTGAACCACTGAAGACTTTGAATTTATTTCAGGAACACGAATGCGCAA 1804
QY 1749 GCTCGCCCGGAGGAGAGAGATCCCGAGATGCTTCATGCGCCCCCAACCATGGAAGGT 1808
DB 1805 ACTTGTGCGAAGAGGCGCAACCACTGAAGGTTTTCATGGCTTCCCAAGGCTTGGACCGT 1864
QY 1809 CTTGACAGATTTATTACAGGTCCCTCGGAGAA 1838
```

[illegible]

Db 1814 ACTTGCTCGAAGGCCAGAAACACCTGACGTTTCATGGCTCCCAAGGCTTGACCGT 1873  
QY 1809 CCGTACAGATATATACAGGTCCTCGAGAA 1838  
Db 1874 GCTGACAGAACTACTACAAATCCTTGAGAA 1903

RESULT 14  
AX281701  
LOCUS AX281701 2646 bp DNA linear PAT 02-NOV-2001  
DEFINITION Sequence 110 from Patent WO0177389.  
ACCESSION AX281701  
VERSION AX281701.1 GI:16608952  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Sellhame,J.J., Porter,G.J.,  
Mikita,T. and Tai,J.  
TITLE Genes expressed in foam cell differentiation  
JOURNAL Patent: WO 0177389-A 110 18-OCT-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1..2646  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 238824.2"  
BASE COUNT 715 a 566 c 634 g 727 t 4 others  
ORIGIN

Query Match 53.7%; Score 990; DB 6; Length 2646;  
Best Local Similarity 71.3%; Pred. No. 8.9e-272;  
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;

QY 9 GATCAAGAAGCAAAAGACGGAGAACACACAGCAAAATCCACCAATCTAGTCTATCAGGCCCA 68  
Db 220 GAGCAATACCCGACAGACTGGGAATGCAGAGAGCAACCAATGTCACTACCAAGCCCA 279  
QY 69 CCATGTGAGCAGGAATTAAGAGAGGGCAAGTGGTGGAAACAGGGGTGGTTCGAGGATG 128  
Db 280 TCATGTGAGCAGGAACAGAGAGGTGAGTGGTGGGACACAGAGGTGGCTTTCGTTGGT 339  
QY 129 TACCGTGTGCTAACAGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188  
Db 340 CACAGTTTGGCTACAGGCTGTCTGAGCGGGAAGAACTACTGTGAGCATGGCCTTGGGA 399  
QY 189 GGAGTACCTTGTCTCCATGCCATCCCTTGTCTTACTCCCTGGATGGGACAAATGTCCGTCA 248  
Db 400 GGAGTACCTGTTGTCTATGCTATCCATCTGCTACACTCTGATGGTGACAAATTCGTCA 459  
QY 249 TGGCCTTAACAGAAATCTCGATCTCTCTGGGACAGAGAGAAATATCCGCGGAT 308  
Db 460 AGGTCTCAATAAATAATCTTGGCTTTAGTCTGAGACAGAGAGAAATGTTCGAGGCT 519  
QY 309 TGGTGGGTGGCTAAGCTGT 368  
Db 520 CGCAGAGTTGCTTAACATGTTTGCAGATGCTGGCTAGTGTGCTACAAAGTTTCATATC 579  
QY 369 TCCATTGCGAAGAGATCGTGAGATGCCGCCAAATATACATGAATCAGAGGGGTGCCATT 428  
Db 580 ACCTTACATCAGGATCGCAACAAATGCAAGGCAAAATTCATGAAGTGCAAGTTTACGGT 639  
QY 429 CTTTGAATATTTGTAGATGCACCTCTAATATTTTGTGAAGCAGAGAGATGAAGGCT 488  
Db 640 TTTTGAAGTATTTTGTAGTCTCCTCTGCTGATGTTTGTGAACAGAGAGATGTCAGAGGACT 699  
QY 489 CTATAAAGGGCCAGAGCTGGGAGATTAAGGATTTACAGTATTTACAGTATTTGATTTCTGATATGA 548  
Db 700 CTACAAAAGCCCGCAGGAGAAATTAAGTTTTCACCTGGATCGATTTCTGAATATGA 759  
QY 549 GAAACCTGAACCTCTCGAGGCTGTGCTTAAACCAAAATTTTCCACAGTAGTGACTGTGT 608

Db 760 AAAGCCAGAGGCCCTGAGTTGGTCTGAAAAACAGACTCCTGTGATGTAAATGACTGTGT 819  
QY 609 CCACCAGGTAGTGGAACTTCTCCAGAGCAGAGAACATTTACCTATATCTATATCAAGA 668  
Db 820 CCAGCAAGTTGTGGAATCTTACAGAAACGGGATATCTTACCTGTGGATGATCTATATGA 879  
QY 669 TATCCACGAACCTCTTTGTGCCGGAACAAACTTTCACCACTGCCAGCTGAGGCTGAAC 728  
Db 880 AGTAAAGAACTATATGTGCCAGAAATTAACCTTCACTTTGGCAAAACAGATGCGGAAC 939  
QY 729 TCTCCCTTCATATCAATTAAGTCTGATCTCCAGTGGGTGCCAGTTTGTAGCAGAG 788  
Db 940 ATTACCAAGCACTGAAATTAATAAGTGGATGATGAGTGGGTGGAGGTTTGGCAGAGG 999  
QY 789 CTGGCCCACTCCCTCAAGGTTTCATCGGGAGAGAGTACTTACAGGTTATCCACTT 848  
Db 1000 TTGGGCAACCCCATTAAGTCTTATGAGAGAGAGGAGTCTTGCAGTGCCTTCATTT 1059  
QY 849 TGACACCTCTAGATGATGGCTGATCAACATGAGCATCCCCCATTTGCTACTGCCGCTC 908  
Db 1060 TGATTTGCTTCTGATGAGGTGCTAATTAACCTTGTCACTACCTATAGTTCTGACTGCGAC 1119  
QY 909 TGCAGAGATAAGACACGGCTGGAGGTGAGCAAGTTTGTCTGACACATGTTGGAGC 968  
Db 1120 TCATGAAGATAAGAGAGAGGTGGAGCGCTGTACAGCATTTTGTCTGTATGATGAGGCGC 1179  
QY 969 GAGGTAGCTATCTTACGAGACGCTGAATTCATGAACACAGAAAGAGAACCTGTTC 1028  
Db 1180 CCGTGTGGCCATCTTCCGAATCCAGATTTTGTGACACAGGAAAGAGGCGCTGTGC 1239  
QY 1029 CCGTGTGGGGACAACTGTACAAACACCCCATATCAAAATGTTGATGGAAGTGG 1088  
Db 1240 CAGACAGTGGGAAACGATGCAAGAACACCCCTATATTAAGATGTTGATGGAACAGG 1299  
QY 1089 GGACTGCTGTGGTGGAGACCTTCAAGTCTGAGGAAATAGATGGAATGATGGCT 1148  
Db 1300 AGATTGGCTGATTGGAGGAGATCTTCAAGTCTTGGATCGAGTTTATTTGGAATGATGTCT 1359  
QY 1149 GGACCAATACCGTCTGACACCTCTGAGAGCTCAAAACAGAAATGTAAGAAATGATGCTGA 1208  
Db 1360 TGATCAGTATCGTCTTACTCTACTGAGCTAAAGCAGAAATTTAAAGATATGATGCTGA 1419  
QY 1209 TGGCGTGTGGATTCAGTTGGCAATCTGTCCACAATGGCCATGCCCTGTGATGCA 1268  
Db 1420 TGTGTCTTTTGCATTTCAACTACGCAACCCAGTGCACATGGACATGCCCTGTGTAATGCA 1479  
QY 1269 GGACACTGCGCGCAGGCTCTAGAGAGGCTCTACAAGCACCCGCTCTCTACTACACCC 1328  
Db 1480 GGATACCCATAAGCAACTCTAGAGAGGCTACCGCGGCCCTCTCTCTCTCTCCACCC 1339  
QY 1329 TCTGGCGGCTGGACCAAGGATGACGATGTGCTCTAGACTGGCGGATGAAGCAGCAGC 1388  
Db 1540 TCTGGGTGGCTGACAAAGGATGACGATGTCTCTTGTGATGTCGCGTATGAAGCAGCATGC 1599  
QY 1389 GGCTGTGCTCGAGGAAGGGTCTGATCCCAAGTCAACCATTTGTTGCCATCTTCCGCTC 1448  
Db 1600 TGCAGTGTGGAGGAAGGAGTCTGAACTCTGAGACGACAGTGTGGCCATCTCCCATC 1659  
QY 1449 TCCCATGTTATATGCTGGGCCCCACAGAGGTCCAGTGGCAGTCCCGGATGATTGC 1508  
Db 1660 TCCCATGATGATGCTGGACCAACTGAGGTCCAGTGGCATTTGACAGAGCAGGATGGTGC 1719  
QY 1509 GGTGTCCAATTTCTACATTTGGGGAGGAGCCCTGCAGGAATCCCCCATCTCTGAAACCAA 1568  
Db 1720 AGAGGCAACTTTTACATTTGTTGACGAGACCCCTGCTGGCATGCCCTCATCCACAAACAGG 1779  
QY 1569 GAAGGATCTGTGAACCCACTCATGCGGCGCAGGTCTTGAGCATGGCCCTGGCCTCAC 1628  
Db 1780 GAAGGATCTTTATGACCAAGTCAATGGTCCCAAGTGTGACAGTGGCCCTGGTTAAT 1839  
QY 1629 CTCTGTGGAATCATTTCCATTTCCGAGTGGCTGCTACAAACAGCCAAAGAACCCATGGA 1688

Db 1840 CACTTGGAAATAGTCTCCCTTTCAGTTGCGAGCTTACAAAGAAAAGAACGCTATGGA 1899  
Qy 1689 CTTCTATGATCCAGCAGGACCAATGAGTTGACTTCATCTCAGGAATCGAATGAGGAA 1748  
Db 1900 CTACTATGACTCTGAACACCATGAGACTTTGAATTTATTCAGAACACCAATGCGGAA 1959  
Qy 1749 GCTCGCCCGGGAAGAGAGATCCCCAGATGGCTTCATGCCCCCAGGAGCATGGAAGGT 1808  
Db 1960 ACTTGCTCGAAGAGGCGAGAACCAACCTGAAGGTTTCATGGCTCCCAAGGCTTGACCGT 2019  
Qy 1809 CTTGACAGATATTACAGGTCCTTGAGAA 1838  
Db 2020 GCTGACAGATATTACAAATCCTTGAGAA 2049

RESULT 15  
BC011392  
LOCUS BC011392 2430 bp mRNA linear PRI 30-JUL-2001  
DEFINITION Homo sapiens, Similar to PAPS synthetase, clone MGC:9898  
ACCESSION IMAGE:3869484, mRNA, complete cds.  
VERSION BC011392  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2430)  
Strausberg, R.  
Direct Submission  
Submitted (25-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590;  
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [villalobebcm.tmc.edu](mailto:villalobebcm.tmc.edu)

Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,

A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,

Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 13 Row: 1 Column: 18

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

#### FEATURES

Location/Qualifiers

1..2430

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="MGC:9898 IMAGE:3869484"

/tissue\_type="Eye, retinoblastoma"

/clone\_lib="NIH-MGC-67"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

21..1832

/codon\_start=1

/product="Similar to PAPS synthetase"

/protein\_id="AAH11392.1"

/db\_xref="GI:15030252"

TTVSMALAEVLVCHGIPCYTLGDNIQGLNKLGFSPDEENVRRIAEVAKLPADA

GLVCITFSIPYDORNNARQIHGASLPFEVFDAPLVHCEORDVGLKFKRAGE

IKGPTGIDSEYEKPEAPLVLKTDSCDNDVCVOOVVLLQERDIPVDASVEVKELV

#### CDS

PENKHLIAKTAETLIPALKINKVDMQWVQVLAEGWATPLNGFREREYLOLHDFCLIL  
DGGVILNSVIPILVATKEDKRGDCTAFALMYERRVAILRNEPFFHRKEERCARQ  
WGTTCKNHPYIKVMWEOCDLIGDLOVLDVYNDGIDLOVRLPTKLFKQKDMNAD  
AVFAFOLRNPVHNGHALLMQDTHKQLBERGYRRPVLLHLLPLGGTCKDDVPLMRMKQ  
HAALVEGVLNPETIVVAIFSPMMIAGPTEQWHRARMVAGANFYILVGRDPAGMPH  
PETKDKLYEFSHGAKVLTMAPLJITLEIVPFRVAAYNKKKKRMDYDSEHHEDEFFTS  
GTRNRKILAREGQKPEGFMAKPTVTEYYKSLKA"

BASE COUNT 688 a 475 c 567 g 700 t

Query Match 53.6%; Score 988.6; DB 9; Length 2430;  
Best Local Similarity 71.8%; Pred. No. 2.2e-271;  
Matches 1294; Conservative 0; Mismatches 509; Indels 0; Gaps 0;

Qy 36 GCAGAAATCCACCAATGTAGTCTATCAGGCGCCACCATGTGAGCAGCAATAAGAGAGGCA 95  
Db 23 GCAGAGACCAACCAATGTACCTACCAAGCCCATCATGTCCAGGACACAGAGAGGTCA 82  
Qy 96 AGTGGTTGGAACAAGGGTGGTTCGAGGAGTATCCGCTGTGGCTTAACAGGTTCTCTGG 155  
Db 83 GGTGGTGGGACACAGAGGTGGTTCGTTGGTTCACAGCTTTGGCTTAACAGGCTTCTCG 142  
Qy 156 TGCTGGAACCAACAGATAAGTTTGGCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCC 215  
Db 143 AGCGGAAAGACTACTGTGAGCATGGCCCTGGAGGAGTACCTGTGCTGATGGTATTTCC 202  
Qy 216 TTGTTACTCCCTGGATGGGACAAATGTCCGCTCATGGCTTAACAGAAATCTCGGATTTCT 275  
Db 203 ATGCTACACTCTGGATGGTGCACATATTCGTCAGGCTCTCAATAAAATCTTGGCTTAG 262  
Qy 276 TCCTGGGACAGAGAGAGAAATATCCGCCGATGTCTGAGTGGCTTAAGCTTTGCTGA 335  
Db 263 TCCTGAAGACAGAGAGAGAAATGTTCGACGATCGCAGAGTTGCTAAACTGTTTGCAGA 322  
Qy 336 TGCTGGTCTGCTGCTATTCACAGCTTTATTTCTCCATTCGCAAGGATCGTGAATGCC 395  
Db 323 TGCTGGCTTAGTGTGCATCAAGTTTCATATCACCTTACATCAGGATCGCAACATGCG 382  
Qy 396 CCGCAAAATACATGAATCAGCAGGCTGCCATTTCTTTGAAATATTTGTAGATGCACTCT 455  
Db 383 AAGCAAAATCATGAAGTGCAGTTTACCGTTTTTTCAGTATTTTGTGATGCTCTCT 442  
Qy 456 AAATATTTGTGAAAGCAGAGAGAGTAAAGGCTCTATAAAAGGGCCAGAGTGGGAGAT 515  
Db 443 GCATGTTGTGAACAGAGGGATGTCAAAGGACTCTACAAAAAGCCGGCAGGAGAAAT 502  
Qy 516 TAAGGATTTACAGGTATTGATTCGATATATGAGAAACCTGAACTCTCAGCGTGTGCT 575  
Db 503 TAAAGGTTTCACTGGGATTCGATATGAAAGCCAGAGGCCCCCTGAGTTGGTGTCT 562  
Qy 576 TAAACCAATTTGTCACAGTGAAGTGTGTCCACCAAGGTAGTGGAACTTTCTCAAGA 635  
Db 563 GAAACACACTCTGTGATGTAATGACTGTGTCCAGCAAGTTGTGGAACTTTCTACAGA 622  
Qy 636 GCAGAACATTCACCTTACTATATAATCAAGATATCCACGAACTCTTTTGGCCGGAAGA 695  
Db 623 ACGGATATTGTACCTGTGGATGATCTTATGAAGTAAAGAACTATATGTGCCAGAGAA 682  
Qy 696 CAAACTTGACCACTCCGAGCTGAGGCTGAACCTCTCCCTTCATTTCAATTAATTAAGT 755  
Db 683 TAAACTTCATTTGGCAAAACAGATGCGGAAACATTTACGACACTGAAAAATTAATAGT 742  
Qy 756 GGATCTCCAGTGGGTCCAGGTTTTCAGCAAGGCTGGGCCACTCCCTCCCAAGGTTTCAT 815  
Db 743 GGATATGCAATGGGTGGTGGAGTTTGGCAGAGGTTGGCAACCCCATTAATGGCTTAT 802  
Qy 816 GCGGAGAGAGGAGTACTTACAGGTATGCACTTTTGACACCTCTGTAGTAGTATGGGTGAT 875  
Db 803 GAGAGAGAGGAGTACTTGCAGTGCCTTCATTTTGAATTTGTCTTCTGATGGAGGTGAT 862  
Qy 876 CAACATGAGCATCCCCATTGTACTGCCGCTCTCTGCAGAGGATAGACACCGCTGGAAG 935  
Db 863 TAACCTGTCTACTACCTATAGTTCTGACTGCGACTCATGAAGATAAAGAGAGGCTGGAC 922



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 14:57:46 ; Search time 445 Seconds  
(without alignments)  
9336.937 Million cell updates/sec

Title: US-09-898-165B-9

Perfect score: 1845

Sequence: 1 atgtcggggaatcaagaagca.....ggtcctggagaagaactaa 1845

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description                 |
|------------|--------|-------------|--------|----|-----------------------------|
| 1          | 1813.6 | 98.3        | 2617   | 21 | AA294206 Human transferase  |
| 2          | 990    | 53.7        | 1875   | 20 | AA84897 Human APS Kinase/A  |
| 3          | 990    | 53.7        | 2507   | 19 | AAV33482 Human disease rela |
| 4          | 990    | 53.7        | 2527   | 20 | AA300717 Human secreted pro |
| 5          | 990    | 53.7        | 2527   | 20 | AA300686 Human secreted pro |
| 6          | 990    | 53.7        | 2646   | 24 | AA594855 Human DNA sequence |
| 7          | 634    | 34.4        | 2463   | 23 | ABL29667 Drosophila melanog |
| c          | 520.8  | 28.2        | 17196  | 23 | ABL29666 Drosophila melanog |
| 8          | 484.8  | 26.3        | 990    | 16 | AAT04849 Human adenosine 5' |

|    |       |      |       |    |                             |
|----|-------|------|-------|----|-----------------------------|
| 10 | 384   | 20.8 | 1695  | 21 | AAC34705 Arabidopsis thalia |
| 11 | 383   | 20.8 | 3684  | 21 | SAMS promoter::ATP          |
| 12 | 378.8 | 20.5 | 401   | 21 | Human colon cancer          |
| 13 | 378.8 | 20.5 | 401   | 21 | Human colon cancer          |
| 14 | 360.8 | 19.6 | 1687  | 21 | Arabidopsis thalia          |
| 15 | 359.8 | 19.5 | 1533  | 21 | Arabidopsis thalia          |
| 16 | 357.6 | 19.4 | 1534  | 21 | Arabidopsis thalia          |
| 17 | 351.2 | 19.0 | 1617  | 21 | Arabidopsis thalia          |
| c  | 315   | 17.1 | 574   | 24 | Gene #360 used to           |
| 18 | 266.4 | 14.4 | 600   | 16 | Human adenosine 5'          |
| 19 | 266.4 | 14.4 | 600   | 16 | Human adenosine 5'          |
| 20 | 233.4 | 12.7 | 435   | 22 | Human breast cance          |
| 21 | 233.4 | 12.7 | 449   | 22 | Human breast cance          |
| c  | 233.4 | 12.7 | 490   | 22 | Human breast cance          |
| 22 | 233.4 | 12.7 | 512   | 22 | Human breast cance          |
| 23 | 233.4 | 12.7 | 512   | 22 | Human breast cance          |
| 24 | 233.4 | 12.7 | 517   | 22 | Human breast cance          |
| 25 | 233.4 | 12.7 | 519   | 22 | Human breast cance          |
| c  | 233.4 | 12.7 | 13418 | 22 | Human immune/haema          |
| 26 | 226   | 12.2 | 2323  | 19 | Human secreted pro          |
| 27 | 206.4 | 11.2 | 692   | 21 | Arabidopsis thalia          |
| 28 | 201.6 | 10.9 | 399   | 24 | Human ovarian anti          |
| 29 | 195.6 | 10.6 | 471   | 22 | Human foetal liver          |
| 30 | 195.6 | 10.6 | 471   | 22 | Human foetal liver          |
| 31 | 195.6 | 10.6 | 471   | 22 | Probe #6431 for ge          |
| 32 | 195.6 | 10.6 | 471   | 22 | Human brain expres          |
| 33 | 195.6 | 10.6 | 471   | 22 | Human bone marrow           |
| 34 | 195.6 | 10.6 | 471   | 22 | Probe #7672 used t          |
| 35 | 195.6 | 10.6 | 471   | 24 | Human genome-deriv          |
| 36 | 195   | 10.6 | 195   | 22 | Human foetal liver          |
| 37 | 195   | 10.6 | 195   | 22 | Probe #16292 for g          |
| 38 | 195   | 10.6 | 195   | 22 | Human brain expres          |
| 39 | 195   | 10.6 | 195   | 22 | Human bone marrow           |
| 40 | 195   | 10.6 | 195   | 22 | Probe #20738 used           |
| 41 | 195   | 10.6 | 195   | 24 | Human genome-deriv          |
| 42 | 169.2 | 9.2  | 1160  | 21 | Saccharomyces carl          |
| 43 | 165   | 8.9  | 222   | 21 | Rat hepatocyte carl         |
| 44 | 157.2 | 8.5  | 935   | 21 | Soybean Adenylisu           |
| 45 | 152   | 8.2  | 1217  | 21 | Corn Adenylisulph           |

ALIGNMENTS

RESULT 1

AA294206

ID AA294206 standard; cDNA; 2617 BP.

XX AC AA294206;

XX DT 19-JUN-2000 (first entry)

XX DE Human transferase TRNSFS-6 CDNA clone 1420940CB1.

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;





|          |   |  |      |
|----------|---|--|------|
| Db       | 1713                                      | TTGAGCATGGCCCTGGCCCTACCTCTGTGGAAATCATTCATTCGAGTGGCTGCCTAC              | 1777 |
| Qy       | 1666                                      | AACAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTC            | 1725 |
| Db       | 1773                                      | AACAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTC            | 1832 |
| Qy       | 1726                                      | ATCTCAGGAACCTCGAATGAGGAAGCTGCGCCCGGAAGGAGAGAATCCCCAGATGGCTTC           | 1785 |
| Db       | 1833                                      | ATCTCAGGAACCTCGAATGAGGAAGCTGCGCCCGGAAGGAGAGAATCCCCAGATGGCTTC           | 1892 |
| Qy       | 1786                                      | ATGGCCCCCAAAAGCATGGGAAGTCCCTGACAGATTATTACAGGTCCTCGGAGAGAAGACTAA        | 1845 |
| Db       | 1893                                      | ATGGCCCCCAAAAGCATGGGAAGTCCCTGACAGATTATTACAGGTCCTCGGAGAGAAGACTAA        | 1952 |
| RESULT 2 |   |  |      |
| AAx84897 |   |  |      |
| ID       | AAx84897 standard; cDNA to mRNA; 1875 BP. |  |      |
| XX       | AC  | AAx84897;  |      |
| XX       | DT  | 24-SEP-1999 (first entry)  |      |
| XX       | XX  | Human APS kinase/ATP sulphurylase gene.                                |      |
| DE       | XX  | APS kinase/ATP sulphurylase; human; PAPS production;                   |      |
| XX       | KW  | 3'-phosphoadenosine 5'-phosphosulphate; ds.                            |      |
| XX       | OS  | Homo sapiens.  |      |
| XX       | XX  | JP11187883-A.  |      |
| PN       | XX  | 13-JUL-1999.   |      |
| XX       | XX  | 26-DEC-1997; 97JP-0360387.   |      |
| PF       | XX  | 26-DEC-1997; 97JP-0360387.   |      |
| XX       | PR  | (NIRA ) UNITIKA LTD.   |      |
| XX       | PA  | (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.                                    |      |
| PA       | XX  | WPI; 1999-451549/38.   |      |
| DR       | XX  | P-PSDB; AAY22349.  |      |
| XX       | XX  | New human-derived APS kinase/ATP sulfurylase gene - useful for         |      |
| PT       | PT  | large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate       |      |
| PT       | PT  | (PAPS)   |      |
| XX       | XX  | Claim 3; Page 7-8; 9pp; Japanese.                                      |      |
| CC       | CC  | This sequence encodes the human-derived APS Kinase/ATP sulphurylase of |      |
| CC       | CC  | the invention. The enzyme may be used to prepare 3'-phosphoadenosine   |      |
| CC       | CC  | 5'-phosphosulphate (PAPS) on a large scale.                            |      |
| XX       | XX  | Sequence 1875 BP; 526 A; 379 C; 499 G; 471 T; 0 other;                 |      |
| SQ       | XX  | Query Match 53.7%; Score 990; DB 20; Length 1875;                      |      |
|          | XX  | Best Local Similarity 71.3%; Pred. No. 5.4e-295;                       |      |
|          | XX  | Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;        |      |
| Qy       | 9   | GATCAAGAGCAAAAGACGGAGAACCCAGCAGAGAATCCACCAATGTAGTCTATCAGGCCCA 68       |      |
| Db       | 39  | GAGCAATAACGCCGACGAACCTGGGGAATGCAGAGAGCAACCAATGTCAACCTACCAAGCCCA 98     |      |
| Qy       | 69  | CCATGTGACAGAGAAATAAGAGAGGGCAAGTGGTTGGAACAAGGGTGGGTTCCGAGGATG 128       |      |
| Db       | 99  | TCATGTGACAGAGAAACAAGAGAGGTCAAGTGGTGGGACCAGAGGTGGCTTCGTGGTTG 158        |      |
| Qy       | 129                                       | TACCGTGTGGCTAACAGGTCTCTCTGGTGTCTGGAAGAAACAACGATAAGTTTTCGCCCTGA 188     |      |
| Db       | 159                                       | CACAGTTTGGCTTACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCCCTTGA 218        |      |
| Qy       | 189                                       | GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGAACAATGTCGGTCA 248       |      |

Db 1299 GGATACCCATAAGCAACTTCTAGAGAGGGGCTACCGGGCCCTGCTCTCTCTCCACCC 1358

Qy 1329 TCTGGGGGCTGGACCAAGATGACGATGTCCTCTAGACTGGCGGATGAAGCAGCAGC 1388

Db 1359 TCTGGTGGCTGGACCAAGATGACGATGTCCTCTAGACTGGCGGATGAAGCAGCAGC 1418

Qy 1389 GGCTGTCTCGAGAGGGTCTGGATCCCAAGCTCAACCTATGTTGCCATCTTCCGTC 1448

Db 1419 TGCAGTGTGGAGGAAGGTTCTGAATCCTGAGACGACAGTGTGGCCATCTTCCCATC 1478

Qy 1449 TCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCACTGCGAGTCCCGGATGATTGC 1508

Db 1479 TCCCATGATGATGCTGGACCAACTGAGTCCAGTGGCACTGAGACACGAGTGGTTGC 1538

Qy 1509 GGGTGGCAATTTCTACATTTGGGGAGGACCTCGAGGAATGCCCATCTCTGAAACCAA 1568

Db 1539 AGGAGCCAACCTTTACATTTGACGAGAGACCTGCTGGCATCCCTCATCCAGAAACAG 1598

Qy 1569 GAAGGATCTGTGAACCCACTCATGCGGCAAGGCTTGGAGTGGCCCTGCGCTCAC 1628

Db 1599 GAAGGATCTTTATGAGCAAGTCACTGTCGCAAGTCTGACGATGGCCCTGCTTAAAT 1658

Qy 1629 CTCTGTGGAATCATTCATTCGAGTGGCTGCTACAAACAAAGCCAAAGCCATGGA 1688

Db 1659 CACTTGGAAATAGTCCCTTTGAGTTGAGCTTACAAAGAAAGAGCGTATGGA 1718

Qy 1689 CTCTATGATCCAGCAAGGACCAATGAGTTTGAATCTATCTAGGAACTCGAATGAGAA 1748

Db 1719 CTACTATGACTCTGAACACCATGAAGACTTTGAATTTATTCAGGACACAGATCGCAA 1778

Qy 1749 GCTGCGCGGAGGAGAGAAATCCCGAGATGGCTTCATGCGCCCAAGCATGGAGGT 1808

Db 1779 ACTGTGTCGAGAGGCGCAAGACCTGAAAGTTCATGGCTCCCAAGGCTTGGACCGT 1838

Qy 1809 CCTGACAGATATTATACAGGTCCTGGAGAA 1838

Db 1839 GCTGACAGATACACAAATCCTTGGAGAA 1868

RESULT 3

AAV33482

ID AAV33482 standard; DNA; 2507 BP.

AC AAV33482;

XX 29-DEC-1998 (first entry)

XX Human disease related nucleotide kinase-2 (DRNK-2) DNA sequence.

XX Human disease related nucleotide kinase-2; DRNK-2; deoxyguanosine kinase; p21ras; cell proliferation; oncogenesis; cancer; gene therapy;

XX Immune disorder; neurological dysfunction; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 37..1911

FT FT /\*tag= a

FT FT /product= "Human DRNK-2"

XX US5817482-A.

XX PD 06-OCT-1998.

XX PF 20-JUN-1997; 97US-0879561.

XX PR 20-JUN-1997; 97US-0879561.

XX PA (INCY-) INCYTE PHARM INC.

XX Bandman O, Corley NC, Guegler KJ, Hawkins PR, Hillman JL;

XX WPI; 1998-556388/47.

DR P-PSDB; AAW70494.

XX Nucleic acids encoding de:oxy:guanosine kinase - useful for

PT recombinant production of the enzyme for treating diseases caused by

PT lack of the enzyme e.g. cancers caused through loss of enzyme

PT function

XX Examples; Fig 2A-2G; 53pp; English.

XX The present sequence represents a human disease related nucleotide

CC kinase-2 (DRNK-2) DNA sequence first identified in Incyte Clone

CC 373887 from the lung tissue cDNA library (JUNGNOT 02). The DRNK-2 DNA

CC sequence is useful for the production of the corresponding recombinant

CC enzyme. The invention provides DRNK enzymes, which are deoxyguanosine

CC kinases, which catalyse the transfer of a terminal phosphate from

CC adenosine triphosphate (ATP) or guanine triphosphate (GTP) to guanosine

CC or guanine in the regulation of cellular levels of GTP and its

CC corresponding nucleoside triphosphate. As GTP levels are known to

CC control the activity of certain oncogenic proteins e.g. p21ras, a protein

CC involved in cell proliferation and oncogenesis, suppression of the enzyme

CC activity causes high ratios of GTP:GDP, promoting oncogenesis.

CC Therefore, diseases (e.g. cancers, immune disorders and neurological

CC dysfunction) related to this lack of activity may be prevented or

CC treated with the recombinant enzyme, or by gene therapy based strategies.

CC Anti-sense constructs of the DRNK encoding nucleic acids may also be

CC used for inhibition of over-expression of the enzyme.

XX Sequence 2507 BP; 704 A; 500 C; 592 G; 710 T; 1 other;

SQ

Query Match 53.7%; Score 990; DB 19; Length 2507;

Best Local Similarity 71.3%; Pred. No. 6.4e-295;

Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;

Qy 9 GATCAAGAAGCAAAAGACGGAGAACCCAGACAGAAATCCACCAATGTAGTCTCTACAGGCCCA 68

Db 75 GACCATATACGCCCAAGAACTGGGGATGCAGACAGACACCAATGTACCTACCAAGCCCA 134

Qy 69 CCATGTGACGAGAAATAAGAGAGGGCAAGTGGTTGGAACAAGGGTGGTTCGAGGATG 128

Db 135 TCATGTGACGAGCAACAAGAGAGGTGAGTGGTGGGACAGAGGTGGCTTCTGCTGGTTG 194

Qy 129 TACCGTGTGGCTAACAGGTCTCTCTGGTCTGGAACCAACCAATAGTTTTCCTGGCA 188

Db 195 CACAGTTTGGCTAACAGGCTTGTCTGGAGCGGGAAGAGACTACTGTGAGCATGGCCCTTGA 254

Qy 189 GGAGTACCTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCCTCA 248

Db 255 GGAGTACCTGGTTTGTCTATGTTTGTATTCATGCTACACTCTGGATGGTGACAAATTCGTCA 314

Qy 249 TGGCCTTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATAATCCCGCGGAT 308

Db 315 AGGTCTCAATAAAATCTTGGCTTTAGTCTCTGAAGACAGAGAGAAGATGTTCCGAGCAT 374

Qy 309 TCGTGAAGTGGCTAACGCTGTTTGGCTGATGCTGTGCTGTGCTGATTCACAGCTTTATTTC 368

Db 375 CCGAGAAAGTTGCTAAACTCTTTGCAGATGCTGGCTTAGTGTGATGACAAAGTTTCATATC 434

Qy 369 TCCATTCCGAAAGGATCGTGAGAAATCCCGCAAAATACATGAATCAGAGGGCTGCCATT 428

Db 435 ACCTTACACTCAGGATCGCAACAATCAAGGCAAAATTCATGAAGTGCAGTTTACCGTT 494

Qy 429 CTTTGAATATTGTTAGATGCACCTCTAAATATTTCTGAAAGCAGAGAGCTAAAGGCCT 488

Db 495 TTTTGAAGTATTGTTGATGCTCTCTGCTGATGTTTGTGAACAGAGGATGTCAGAGGACT 554

Qy 489 CTATAAAGGGCCAGAGCTGGGAGATTAAGAGTTTACAGGTTTACATTCGATTATGA 548

Db 555 CTACAAAAGAGCCCGGAGAGAAATTAAGGTTTCACTGGGATCGATTTCTGAATATGA 614

Qy 549 GAAACCTGAAACTCTCTGAGCGTGTGCTTAAACCAATTTGTCCACAGCTGAGTGTGTGT 608

Db 615 AAGCCAGAGGCCCTGAGTGTGCTGAAACAGAGACTCTGTGATGATTAATGACTGTGT 674

QY 609 CCACAGTAGTGAACCTCTGCAAGAGCAGAACTTGTACCCCTATATCTAATAAACA 668  
Db 675 CCAGCAAGTTGTGAACTTCTACAGGAACGGATATTGTACCTGTGGATGACTTATGA 734  
QY 669 TATCCAGAACTCTTTGTGCGGGAAGAAACAACACTTGCACACGCTCCGAGCTGAAC 728  
Db 735 AGTAAAGAACTATGTGCGGGAAGAAATAAATCTTATTTGGCAAAAACAGATCGGAAAC 794  
QY 729 TCTCCCTTCTATTATCAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTTTGACGGAAG 788  
Db 795 ATTACCACTGAAATTAATAAGTGGATACGAGTGGGTGAGGTTTTGGCAGAGG 854  
QY 789 GTGGGCCACTCCCTCAAGGTTTTATCGCGGAGAGGAGTACTTACAGGTTATGCACTT 848  
Db 855 TTGGGCAACCCATTGAATGGCTTTATGAGAGAGGAGGAGTACTTGCAGTGCCTTCATT 914  
QY 849 TGACACCTCTAGATGCGCTGATCAACATGAGCATCCCATTTGACTGCGCTCTC 908  
Db 915 TGATTGCTCTGATGAGGTGCTAATTAAGTGTCTAGTACTATAGTTCTGACTGGCAG 974  
QY 909 TGCAGAGGATAAGACAGGCTGGAAGGCTGAGCAAGTTTGTCTGCGCACATGGTGGACG 968  
Db 975 TCATGAGATAAGAGAGGCTGGACGCTGTACAGCATTTGCTCTGATGATGAGGCGC 1034  
QY 969 GAGGAGTACTTATGAGAGAGCTGAATCTATGACACAGAAAGAGGAGGAGCTGTC 1028  
Db 1035 CCGTGTGGCCATTCTTCCCAATCCAGAGTTTTTTGAGCACAGGAAAGAGGAGCTGTC 1094  
QY 1029 CCGTGTGTGGGGACACATGTACAAACACCCCATATCAAAATGTTGATGGAAGTGG 1088  
Db 1095 CAGACAGTGGGGAACGACATGCAAGACCCCATATTAAGATGTTGATGGAACAAAG 1154  
QY 1089 GGACTGGCTGGTGTGGAGACCTTCAGGTGCTGGAGAAATAAGATGGAATGATGGCT 1148  
Db 1155 AGATTGGCTGATTGGAGAGATCTCAAGTCTGGATCGAGCTTATTGGAATGATGCT 1214  
QY 1149 GGACCAATACCGTCTGACACCTCTGGAGCTCAACACAGAAATGTAAAGAAATGAATGCTGA 1208  
Db 1215 TGATCAGTATGCTTACTCTCTAGCTAAAGCAGAAATTTAAAGATATGAATGCTGA 1274  
QY 1209 TGGCGTGTTCATTTCAGTTGGCAATCCTGTCACAAATGCCATGCCCTGTTGATGCA 1268  
Db 1275 TGCTGTCTTGCATTTCACTACCAACCCAGTGCACATGGATGCGCTGTTAATGCA 1334  
QY 1269 GGACACCTGCGCAGGCTCTTAGAGAGGGGTACAAGCACCCGCTCTCTCTACTACACCC 1328  
Db 1335 GGATACCCATAAGCAACTTCTAGAGAGGGGTACCGCGCCCTGTCTCTCTCTCCACCC 1394  
QY 1329 TCTGGCGGCTGGACCAAGGATGAGATGTCCTCTAGACTGGCGGATGAAGCAGCAGC 1388  
Db 1395 TCTGGGTGGCTGGCAAGGATGACGATGTTCTCTTGTATGTGGCGTATGAAGCAGCATGC 1454  
QY 1389 GGCTGTGCTGAGGAAGGGTCTCGGATCCCAAGTCAACCAATGTTGCGCATCTTCCGTC 1448  
Db 1455 TGCAGTGTGAGGAAGAGTCTGAATCTTGAGACGACAGTGGTGGCGCATCTTCCCATC 1514  
QY 1449 TCCCATGTTATGCTGSCCCACACAGGTCAGTGGCAGCTCCAGGTCGCCGATGATTC 1508  
Db 1515 TCCCATGATGATGCTGGACCAACTGAGGTCCAGTGGCATTTGCAGAGCACGATGTTGC 1574  
QY 1509 GGGTGCCAAATTTACATTTGGGGAGGAGCCCTGACAGGAATGCCCATCTCTGAAACCA 1568  
Db 1575 AGGAGCCCAACTTTTACATTTGTGGACGAGACCCCTGCTGGCATGCTCTCAGAAACAGG 1634  
QY 1569 GAAGGATCTGATGAACCCACTCATGCGGCAAGGCTTGGAGCATGGCCCTGCGCTCAC 1628  
Db 1635 GAAGGATCTTATGAGCAAGTCAATGTTGGCAAGTGTGACGATGGCCCTGTTTAAT 1694  
QY 1629 CTCTGTGGAATCATCTCCAGTGGGTGCTTACAAACAAAGCCAAAGGAGCATGGA 1688  
Db 1695 CACTTTGGAATAGTTCCCTTTCAGTTGAGCTTACAAAGAAAGAGGAGCTATGGA 1754  
QY 1689 CTTCTATGATCCAGAGGCAACATGATGTTGACTTCTATCTCAGGAACCTCGAATGAGGAA 1748

Db 1755 CTACTATGACTCTGAACACCATGAAGACTTTTGAATTTATTTTCAGGAACACCAATGCGCAA 1814  
QY 1749 GCTCGCCGGGAAGAGAGAAATCCCCAGATGGCTTCATGCGCCCAAGCATGAAGGT 1808  
Db 1815 ACTTGTCTGGAGAGGCCAGAAACCACTTGAAGGTTTCATGCTCCCAAGGCTTGGACCGT 1874  
QY 1809 CCTGACAGATTTATTACAGGTCCTCGGAGAA 1838  
Db 1875 GCTCAGCAATACTACAAATCCTTGGAGAA 1904  
RESULT 4  
AX00717  
ID AX00717 standard; DNA; 2527 BP.  
XX AX00717;  
XX AX00717;  
DT 25-MAR-1999 (first entry)  
XX Human secreted protein gene 76 clone HOSFD58.  
DE Human; secreted protein; fusion protein; gene therapy; protein therapy;  
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX Homo sapiens.  
OS  
PN W09842738-A1.  
XX 01-OCT-1998.  
PD  
XX 19-MAR-1998; 98WO-US05311.  
XX 30-MAY-1997; 97US-0050937.  
PR 21-MAR-1997; 97US-0041276.  
PR 21-MAR-1997; 97US-0041277.  
PR 21-MAR-1997; 97US-0041281.  
PR 21-MAR-1997; 97US-0042344.  
PR 30-MAY-1997; 97US-0048069.  
PR 30-MAY-1997; 97US-0048094.  
PR 30-MAY-1997; 97US-0048095.  
PR 30-MAY-1997; 97US-0048096.  
PR 30-MAY-1997; 97US-0048099.  
PR 30-MAY-1997; 97US-0048131.  
PR 30-MAY-1997; 97US-0048135.  
PR 30-MAY-1997; 97US-0048154.  
PR 30-MAY-1997; 97US-0048160.  
PR 30-MAY-1997; 97US-0048186.  
PR 30-MAY-1997; 97US-0048187.  
PR 30-MAY-1997; 97US-0048188.  
PR 30-MAY-1997; 97US-0048350.  
PR 30-MAY-1997; 97US-0048351.  
PR 30-MAY-1997; 97US-0048352.  
PR 30-MAY-1997; 97US-0048355.  
PR 05-AUG-1997; 97US-0054804.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Brewer LA, Duan R, Ferrie AM, Florence KA;  
PI Greene JM, Hu JS, Lafleur DW, Moore PA, NI J, Olsen HS;  
PI Rosen CA, Ruben SM, Shi Y, Young P;  
XX WPI; 1999-070066/06.  
DR P-PSDB; RAN67913.  
XX  
PT New isolated human genes and the secreted polypeptides they encode -  
useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders

PS Claim 1; Page 265-267; 385pp; English.

XX This sequence represents a nucleic acid molecule which encodes a secreted human protein. The gene number, and the clone it is derived from, are detailed in the descriptor line. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAX00602) for increasing the stability of the fused protein as compared to the human protein only.

CC The invention relates to 87 novel genes and their fragments (nucleic acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 87 polynucleotides, based on which tissues they are most highly expressed in (see AAX00611 for described uses).

XX Sequence 2527 BP; 706 A; 508 C; 597 G; 716 T; 0 other;

Query Match 53.78; Score 990; DB 20; Length 2527;

Best Local Similarity 71.38; Pred. No. 6.4e-295;

Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;

|    |     |  |     |
|----|-----|--|-----|
| QY | 9   | GATCAAGAACAAAGACGGAGAACACAGAGAAATCCACCAATGTAGTCTATCAGGCCCA     | 68  |
| DB | 94  | GAGCAATAACCGCGAGAACTGGGGAATGCAGAGAGCAACCAATGTCACTACCAAGCCCA    | 153 |
| QY | 69  | CCATGTGAGCAGGATTAAGAGAGGGCAAGTGGTGGAAACAAAGGGTTCGAGGATG        | 128 |
| DB | 154 | TCATGTGAGCAGGAACAAAGAGAGTCAAGTGGTGGGACAGAGGTGGCTTCGTGGTGG      | 213 |
| QY | 129 | TACGCTGTGCTAACAGTCTCTCTGTGCTGTGAAACAAACAGTAAGTTTGCCTTGA        | 188 |
| DB | 214 | CACAGTTGGCTAACAGGCTTCTGTGAGCGGGAAGACTACTGTGAGCATGGCTTGA        | 273 |
| QY | 189 | GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGCCGTCA     | 248 |
| DB | 274 | GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGCCGTCA     | 333 |
| QY | 249 | TGGCCTTAACAGAAATCCGATCTCTCTGTGCTGTGAAACAAACAGTAAGTTTGCCTTGA    | 308 |
| DB | 334 | AGGTCTCAATAAATACTTGGCTTTAGTCTGTAAGACAGAGAGAAATTTCCAGCGCAT      | 363 |
| QY | 309 | TGCTGAGGTGGCTAAGCTGTTTGTGCTGATGCTGTGCTGTGCTGCAATACCAAGCTTTATTC | 368 |
| DB | 394 | CGCAGAGTTGCTAAACTGTTTCAGATGCTGGCTAGTGTGCTACACAGTTTCATATC       | 433 |
| QY | 369 | TCCATTGCGAAAGATGCTGAGAAATGCCGCAAAATACATGAATCAGCAGGGTGCAT       | 428 |
| DB | 454 | ACCTTACACTCAGGATCGCAACAAATTAAGGTTTCACTGGGATCGATTCTCAATATGA     | 513 |
| QY | 429 | CTTTGAAATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGAGACGCTAAAGGCT       | 488 |
| DB | 514 | TTTGTGAAGTATTTGTGATGCTCCTCTGCAATGTTGTGAACAGAGGGATGTCAAAGGACT   | 573 |
| QY | 489 | CTATAAAGGGCCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATTCTGATTATGA      | 548 |
| DB | 574 | CTACAAAAGCCCGGCGAGAGAAATTAAGGTTTCACTGGGATCGATTCTCAATATGA       | 633 |
| QY | 549 | GAACCTGAAATCTCTGAGGCTGTGCTTAAACCAATTTGTCCACAGTGAAGTGTGT        | 608 |
| DB | 634 | AAAGCCAGAGGGCCCTGAGTTGTGTGCTGTAACACAGACTCCTGTGATGTAATGACTGTGT  | 693 |
| QY | 609 | CCACAGGTAGTGGAACTCTCGAAGACAGAGACATTTGTACCCCTATATAATCAAGA       | 668 |
| DB | 694 | CCAGCAGTTTGTGGAATCTTACAGGAACGGGATATTTGTACCTGTGGATGCTATGA       | 753 |
| QY | 669 | TATCCAGCAACTCTTTGTCCGGAACCAAACTTGTACCCAGCTCCGAGCTGAGGCTGAAAC   | 728 |
| DB | 754 | AGTAAAGAACTATATGTGCCAGAAATAAATTCATTTGTCGCAAAACAGATCGCGAAC      | 813 |

|    |      |  |      |
|----|------|--|------|
| QY | 729  | TCCTCCCTTCATTATCAATTAAGCTGGATCCAGTGGGTGCTCCAGGTGTTTGGCGAAGG      | 788  |
| DB | 814  | ATTACACAGCACTGNAATTAATAAGTGGATATGCCAGTGGGTGAGGTTTGGCAGAAGG       | 873  |
| QY | 789  | CTGGGCACTCCCTCAAGGTTTCATGCGGAGAGAGTACTTACAGGTTATGCACTT           | 848  |
| DB | 874  | TTGGCAACCCCATTTGAATGGCTTTATGAGAGAGAGGAGTACTTGCAGTGCCTTCATTT      | 933  |
| QY | 849  | TGACACCCCTCTAGATGATCGGCTGATCAACATGAGCATCCCCATTGTACTGCCCTCTC      | 908  |
| DB | 934  | TGATTGTCTTCTGGATGGAGGTGTCAATTAAGTGTTCAGTACCTATAGTTCTGACTGCGAC    | 993  |
| QY | 909  | TGCAGAGGATAAGACACAGCTGGAAGGCTGCAGCAAGTTTGTCTGCACATGTTGGACG       | 968  |
| DB | 994  | TCATGAGATTAAGAGAGGCTGCAGCGCTGTACAGCATTTTGTCTGATGTATGAGGGCCG      | 1053 |
| QY | 969  | GAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGAGCGCTGTC        | 1028 |
| DB | 1054 | CCGTGTGGCCATTTCTTGGCAATCCAGAGTTTTTGTGACACAGGAAAGAGGCGCTGTC       | 1113 |
| QY | 1029 | CCGTGTTGGGGACACATGTCAAAACACACCCCATATCAAAATGTTGATGAAAGTGG         | 1088 |
| DB | 1114 | CAGACAGTGGGAAACGACATGCAAGAACCCCTTATTTAAGATGGTGTGAGCAAGG          | 1173 |
| QY | 1089 | GGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATTAAGATGGAATGATGGCT      | 1148 |
| DB | 1174 | AGATTGGCTGATTGGAGGAGATCTCAAGTCTTGGATCGAGTTTATTGGAATGATGCTCT      | 1233 |
| QY | 1149 | GGACCAATACCTGCTGACACCTCTGGAGCTCAACACAGAAATGTAAGAAATGAATGCTGA     | 1208 |
| DB | 1234 | TGATCAGTATCGTCTTACTCTACTAGCTTAAAGCAGAAATTTAAAGATGATGAATGCTGA     | 1293 |
| QY | 1209 | TGCGGTGTTTGGCAATCCAGTTGCGCAATCTGCTCCCAATGGCCCTGTTGTGATGCA        | 1268 |
| DB | 1294 | TGCTGCTTTGCAATTTCACTAGCAACCCAGTGCACATGGACATGCGCTGTTAATGCA        | 1353 |
| QY | 1269 | GGACACTGCGCGAGGCTCTTAGAGAGGGCTTACAAGCACCCGGTCTCTTACTACACC        | 1328 |
| DB | 1354 | GGATACCCATAAGCAACTTCTAGAGAGGGCTACCAGCGCCCTGCTCTCTCTCCACCC        | 1413 |
| QY | 1329 | TCCTGGGCGGCTGACCAAGGATCAGCATGTCCTCTAGACTGGCGGATGAAGCAGACGC       | 1388 |
| DB | 1414 | TCCTGGTGGCTGACAAAGGATGACCATGTTCTTGTATGTCGCTTAAGCAGCATGC          | 1473 |
| QY | 1389 | GGCTGTGCTCGAGAGGGGCTCTGGATCCCAAGTCAACCATTTGTTGCCATCTTTCGCTC      | 1448 |
| DB | 1474 | TGCAGTGTTCGAGAGAGGAGTCTGATCCTGAGACGACAGTGTGGCCATCTTCCATC         | 1533 |
| QY | 1449 | TCCCATGTTATATGCTGGGCCCCACAGAGTCCAGTGGCACTGCGAGTCCCGGATGATGC      | 1508 |
| DB | 1534 | TCCCATGATGATGCTGGACCAACTAGGTGCCAGTGGCATTTGCAGACGCGATGTTGC        | 1593 |
| QY | 1509 | GGGTGCCAATTTCTACATTTGGGAGGGACCTCGAGGAATGCCCATCTCTCAAAACCA        | 1568 |
| DB | 1594 | AGGAGCCAACTTTTACATTTTGGACGAGACCTCTGCTGGCATGCCCTCATCCAGAAACAGG    | 1653 |
| QY | 1569 | GAGGATCTGTATGAACCCACTCATGGGGCAAGGCTTTGAGCATGCGCCCTGGCCCTCAC      | 1628 |
| DB | 1654 | GAGGATCTTTATGAGCCCAAGTCTGTCGCAAGTGTGACCATGCGCCCTGGTTTAAAT        | 1713 |
| QY | 1629 | CTCTGTGGAATCATTTCCATTCGAGTGGCTGCTTACACAAAGCCCAAAAGCCCATGGA       | 1688 |
| DB | 1714 | CACCTTTGGAATAGTTTCCCTTTCAGTTCAGTTCAGCTTACAACAGAAAAGAGCGTATGA     | 1773 |
| QY | 1689 | CTTCTATGATCCAGCAAGGCACAAATGAGTTTGTACTTCTCATGAGAACTGGAATGAGAA     | 1748 |
| DB | 1774 | CTACTATGACTTGAACACCATGAACTTTGAATTTATTTTACAGAACCAAGATGCGCAA       | 1833 |
| QY | 1749 | GCTCGCGCGGAGGAGAGAAATCCCGAGATGGCTTTCATGGCCCCCAAGCATGGAAGGT       | 1808 |
| DB | 1834 | ACTTGTCTCGAGAGGCGCAGAAACCACTTGAAGGTTTTTCATGGCTTCCCAAGGCTTGGACCGT | 1893 |



Db 874 TTGGCAACCCCATTAAGTGGCTTTATGAGAGAGAGGGAGTACTTCCAGTGGCTTCATTT 933  
 QY 849 TGACACCTCTGCTAGATGGCTGATCAACATGAGCATGCCCATTTGCTAGTCCCTCTC 908  
 Db 934 TGATTCTCTCTGGATGGAGGTGTCATTAACCTTGTCTAGTACCTATATGTTCTGACTCGAC 993  
 QY 909 TGCAGAGGATAAGACACGGCTGGAAGGGTGACAAAGTTTGTCTCTGGCACATGGTGGAGC 968  
 Db 994 TCATGAAGATAAAGAGAGGCTGGACGGCTGTACAGCATTTGCTCTGATGTATGAGGGCG 1053  
 QY 969 GAGGTAGCTATCTTACGAGACGCTCAATCTATGAACACAGAAAGAGAGCGCTGTTTC 1028  
 Db 1054 CCGTGGCCCATTTCTCGCAATCCAGAGTTTTTGGACACAGAAAGAGAGCGCTGTGC 1113  
 QY 1029 CCGTGTGGGGGACAAACATGTACAAACACACCCCATATCAAAATGGTGTGAAAGTGG 1088  
 Db 1114 CAGACAGTGGGAACGACATGCAAGAACCCCTATATTAAGATGGTGTGAAACAGG 1173  
 QY 1089 GACTGGCTGGTGGAGACCTTCCAGGTGTGGAGAAATAAAGATGAATGAGGT 1148  
 Db 1174 AGATTGGCTGAATGGAGAGATCTTCAAGTCTTGGATCGAGTTTATTTGAATGATGCTCT 1233  
 QY 1149 GGACCAATACCGCTCTCACACCTCTGAGCTCAACACAGAAATGTAAGAAATGAATCTCA 1208  
 Db 1234 TGATCAGTATGCTTACTCTCTACTGAGCTAAGCAGAAATTTAAAGATATGAATCTGA 1293  
 QY 1209 TGGGTGTTTGGCATTCAGTGGCAATCTGTCTCCCAATGGCCATGCCCTGTGGATGCA 1268  
 Db 1294 TGCTGTCTTGGCATTTCAACTAGCAACCCAGTGCACATGGACATGCCCTGTATATGCA 1353  
 QY 1269 GACACCTCGGAGGCTCTAGAGAGGGCTTACAGACCCGGTCTCTCTACTACACCC 1328  
 Db 1354 GGATACCCCAAGCAACTTCTAGAGAGGGCTACCGCGCCCTGTCTCTCTCTCCACCC 1413  
 QY 1329 TGTGGCGCTGGACCAAGATGACATGCTGTCTGAGACTGCGGATGAAGCAGCACGC 1388  
 Db 1414 TCTGGTGGCTGGCAAGAGATGACATGTTCTTGTATGCTGGGTATGAGCAGATGC 1473  
 QY 1389 GCGTGTCTGGAGAGGGTCTGTGATCCCAAGTCAACCATTTGTGCCATCTTTCCTGTC 1448  
 Db 1474 TGCAGTGTGGAGAGAGGATTTCTGAATCTGAGAGCAGAGTGTGGCCATCTTCCATC 1533  
 QY 1449 TCCTATGTTATGCTGGCCACACAGAGCTCCAGTGGCACTGCAGTCCCGGATGATGC 1508  
 Db 1534 TCCATGATGTATGCTGGACCACTGAGGTCCAGTGGCATTTGCAGAGCAGCGATGTTGC 1593  
 QY 1509 GGTGTCCAATTTCTACATTTGGGGAGGACCTTCAGGAATGCCCATCTCTGAAACCAA 1568  
 Db 1594 AGGAGCAACTTTTACATTTGTTGGAGCAGACCTGCTGGCATGCTCATCCAGAACAGG 1653  
 QY 1569 GAAGATCTGTATGAACCACTCATGGGGCAAGGTCTTTGAGCATGGCCCTGGCCCTAC 1628  
 Db 1654 GAAGGATCTTTATGAGCCAAAGTCATGGTCCCAAGTGTCTGAGCATGGCCCTGGTTAAT 1713  
 QY 1629 CTCTGTGAATCATTTCCATTTCCGAGTGGCTGCTTACACAAAGCCCAAAAGCCATGCA 1688  
 Db 1714 CACTTTGGAATAGTTTCCCTTTTCGAGTTCAGCTTACACAAAGAAAGAGCGTATGGA 1773  
 QY 1689 CTCTATGATCCAGCAAGCACAATGAGTTTTCATCTCTCAGGAACCTCGAATGAGGAA 1748  
 Db 1774 CTATATGACTCTGACACCATCATGAGACTTTGAATTTATTTAGGAACACGAATCGCAA 1833  
 QY 1749 GCTGCCCCGGAGGAGAGAAATCCCCAGATGGCTTTCATGGCCCCCAAGATGGAAGT 1808  
 Db 1834 ACTTGTCTGAGAGGGCCAAACACCTGAAGGTTTTCATGGTCTCCCAAGGCTTGGACCGT 1893  
 QY 1809 CTTGACAGATTTATCAGGTCTCTGGAGAA 1838  
 Db 1894 GCTGACAGAATACTACAAATCTTTGGAGAA 1923

ID AAS94855 standard; DNA; 2646 BP.  
 XX AAS94855;  
 AC  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human DNA sequence #110 expressed during foam cell differentiation.  
 XX  
 KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200177389-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 04-APR-2001; 2001WO-US11128.  
 XX  
 PR 05-APR-2000; 2000US-195106P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Shiffman D, Somogyi R, Lawn R, Sellhamer JJ, Porter GJ, Mikita T;  
 PI Tai J;  
 XX  
 DR WPI; 2002-010925/01.  
 XX  
 CC Composition useful for diagnosis of conditions, disorders or diseases  
 CC associated with atherosclerosis, comprises several polynucleotides that  
 CC are differentially expressed in foam cell development -  
 XX  
 CC Claim 1; Page 162-163; 315pp: English.  
 XX  
 CC The present invention relates to the isolation of human polynucleotide  
 CC sequences that are differentially expressed during foam cell  
 CC differentiation. The polynucleotide sequences of the invention or a  
 CC composition comprising these polynucleotides are useful as a high  
 CC throughput method for detecting altered expression of one or more  
 CC polynucleotides in a sample. The polynucleotides can be used in the  
 CC diagnosis of disorders associated with foam cell development such as  
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
 CC coronary artery disease. The polynucleotide sequences can also be used  
 CC as PCR primers and probes. The polynucleotides of the invention are also  
 CC useful in gene therapy. AAS94746-AAS95021 represent the human  
 CC polynucleotide sequences of the invention which are differentially  
 CC expressed during foam cell differentiation.  
 XX  
 SQ Sequence 2646 BP; 715 A; 566 C; 634 G; 727 T; 4 other;  
 Query Match 53.7%; Score 990; DB 24; Length 2646;  
 Best Local Similarity 71.3%; Pred. No. 6.6e-295;  
 Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;  
 QY 9 GATCAGAGACAAAGACGAGAGGCAAGTGGTGAACAAGGGTGGTTCGAGGATG 68  
 Db 220 GAGCAATAACGCGCAACTGGGGAATGCAGAGAGCAACCAATGTCCATACCAGCCCA 279  
 QY 69 CCATGTGAGCAGGAATAAGAGAGGCAAGTGGTGAACAAGGGTGGTTCGAGGATG 128  
 Db 280 TCATGTGAGCAGGAACAAGAGAGGTCAGGTGGTGGGACGAGGTCGTCGTCGTTG 339  
 QY 129 TACCGTGTGGCTAACAGGTCTCTCTGGTCTCGAAGAAACAGATAAAGTTTGGCCCTGGA 188  
 Db 340 CACAGTTTGGCTTAACAGGCTTGTCTGAGCGGGAAGACTACTGTGAGCATGGCTTGA 399  
 QY 189 GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAATGTCGCTCA 248  
 Db 400 GGAGTACCTTGGTGTTCATGTGGTATTCCTACTCTGAGTGGTGAACAATATTCGTCA 459  
 QY 249 TGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATATCGCCGGAT 308  
 Db 460 AGGTCTCAATAAATCTTGGCTTTAGTCTCTGAAGACAGAGAGAGAAATGTTGCGGCAT 519





CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 2463 BP; 579 A; 667 C; 685 G; 532 T; 0 other;

|                       |              |                     |                 |              |
|-----------------------|--------------|---------------------|-----------------|--------------|
| Query Match           | 34.4%        | Score 634;          | DB 23;          | Length 2463; |
| Best Local Similarity | 60.7%;       | Pred. No. 7.8e-185; |                 |              |
| Matches 1124;         | Conservative | 0;                  | Mismatches 695; | Indels 33;   |
|                       |              |                     |                 | Gaps 4;      |

|      |    |   |      |
|------|----|---|------|
| 27   | QY | GGAGAACCGACGACAGAAATCCACCAATGTAGTCTTATCAGGCCCCACCATGTGAGCAGGAATAA | 86   |
| 297  | Db | GCAGTGCCTACAGCTGGCGAGCAAGTGTGACCGGACGAAAAGCACCATGTCCACGAGGAGAC    | 356  |
| 87   | QY | GAGAGGGCAAGTGTTGGGAACAAGGGGTGGTTCGAGAGATGTACCGTGTGGCTTAACAGG      | 146  |
| 357  | Db | GCGCGCAAGAATTTGGGACTCTGTCTGAGGATTTCCGAGGATGCACGGTGTGGCTGACTGG     | 416  |
| 147  | QY | TCCTCTGTGTGCTGGAAAAACAACAGATAAGTTTTCGCTCGGAGAGTACCTGTCTCCCA       | 206  |
| 417  | Db | CCTCAGTGGAGCTGGAAAGACCTCGATTGCTTTCGAACCTGGAGGCGGTACCTGTGATCCCG    | 476  |
| 207  | QY | TGCACTCCCTTTGTTACTTCCTCGATGGGACAATGTCCGTGATGGCTTAACAGAAATCT       | 266  |
| 477  | Db | GGCATTTCGGGCTACGGTTGGACGCGCAACATTCGCAACCGGACTGAACAGAAGACCT        | 536  |
| 267  | QY | CGGATTCCTCTCGGGACAGAGAGGAAAAATATCCCGCCGATTTCTGTGAGTGGCTTAAGCT     | 326  |
| 537  | Db | GGGCTTTCAGCCCGCGATCGCGAGGAAAAACATTCGCGCGGTGGCGAGGTGGCCAAAGTT      | 596  |
| 327  | QY | GTTTGCTGATGCTGTCTGGTCTGCAATTACCAAGCTTTATTTCTCCATTCCTCCAAAGGATCG   | 386  |
| 597  | Db | GTTTGGCGACGCGGAGTGGTGCCATCTCGAGCTTCGTTTTCGCCCTTCGCCACGATCG        | 656  |
| 387  | QY | TGAGAATGCCGCGAAAAATACATGAATCAGCAGGGCTGCCATCTTTTGAATAATTTGTAGA     | 446  |
| 657  | Db | TGAGATGGCGCGCAAGATCCACAAGGATCGGGTCTGAAGTTCTATGAGATATTCGTGGA       | 716  |
| 447  | QY | TGCACCTCTAATAATTTGTGAAGCAGACGACTAAAGGCCTCTATAAAGGGCGCAGAGC        | 506  |
| 717  | Db | CACACGCTGGATGTCTGCGAAAACCGCGGAGCTGAAGGGTCTATACAAGAAGCGCGCGA       | 776  |
| 507  | QY | TGGGAGATTAAGGATTTACAGGATTTGATCTGTGATTTATGAGAAACCTCAAACTCCTGA      | 566  |
| 777  | Db | GGCGTAAATCAAGGATTCAGGGCATTACCGCAGGAGTACGACGACCCAGATTCGCGGA        | 836  |
| 567  | QY | GGGTGTGCTTTAAACCAATTTGTCCACAGTGAAGTACTGTGCCACAGGTAGTGGAACT        | 626  |
| 837  | Db | GCTGTGGTCAACACGACGGCTACACGGTGCAGGCTGCAGGCTCCACTCAGAAGCTGTGACGCT   | 896  |
| 627  | QY | TCGCAAGACGACAACTGTACCCCTATACTATAATCAAGATATCC---ACGAACCTCT         | 682  |
| 897  | Db | GCTCGAACAGGAGGGGCAATTATCCGCGCTGCTGCGCAGCTAGACCTGTGTCGGGAGCT       | 956  |
| 683  | QY | TTGTGCGGAAAAACAACATT-----GACCACGTCCTGAGCTGAGGCTGAACACTCTCCCTTC    | 737  |
| 957  | Db | TTATCCAGCAATCCATTGCCACGAGGACACTGCGTTCAGGAGCGCGAGTCGTGCAAGC        | 1016 |
| 738  | QY | ATTATCAATTACTAAGCTGGATCTCAGTGGGTCCAGGTTTTCAGGGAAGGCTTGGGCCAC      | 797  |
| 1017 | Db | CATFCGAATTCAGCACCGTGGAGCTGCAGTGGGTGCAAGTACTCGGCCGAGGGATGGGCATA    | 1076 |
| 798  | QY | TCCCTCTCAAAGTTTTCATCGGGAGAGAAGTAGTACTTACAGGTTATGCACCTTTCCACAC---  | 854  |
| 1077 | Db | TCCACTGCGGGCTTCATGCGCGAGGATGNGTACTTCAGACGCTGCATCTCAACACT          | 1136 |
| 855  | QY | -----CCTGCTAGATGATGGCGTGATCAACATGAGCATCCCATTTGACTGCCCGT           | 905  |
| 1137 | Db | TCAGAGCGGCATGGACGGTTTCTTATCGGAGAACCACTCGGTGCCCATTTGTTTGAAGCG      | 1196 |





|          |   |   |      |
|----------|---|---|------|
| Qy       | 1610  | GCATGGCCCTGGCCTCACC-----  | 1629 |
|          |   |   |      |
| Db       | 1615  | AGATGGCCCGGAGCTAGACAGCATGGAGGTAAGGTACACTACCTAATCGCCAAATC        | 1556 |
| Qy       | 1630  | -----TCTGTGGAATCATTCATTCCTCGAGTGGCTGCCTAC                       | 1665 |
| Db       | 1555  | TAGCCGNACTAAATCGATATTTCTTTTAGATCTCGCCCTTCGTGTGCTGCCTAT          | 1496 |
| Qy       | 1666  | AACAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAGGCAACAATGAGTTTGACTTC      | 1725 |
| Db       | 1495  | GACAAGAGCGCCAGTAGAATGGCTTCTTCGAGCCCAAGAGGAAGGATGAGTTTGAGTTC     | 1436 |
| Qy       | 1726  | ATCTCAGGAAGCTCGAATGAGGAAGCTCGCCGGGAAGGAGAGAAATCCCCAGATGGCTTC    | 1785 |
| Db       | 1435  | ATCTCGGGGAACCAAGATGGCCACCTGGCCAAAACGGAGCCAGTCCGCCGATGGCTTC      | 1376 |
| Qy       | 1786  | ATGGCCCCCAAGCATGGGAAGGTCCCTGACAGATTATTACAGGTCCCTGGAGAGAAGAACTAA | 1845 |
| Db       | 1375  | ATGGAAACCGAGGCTTGGCCCATCTGGCCACCTACTACCAGAACCCTGCCGAGTCGTAA     | 1316 |
| RESULT 9 |   |   |      |
| AAT04849 |   |   |      |
| ID       | AAT04849 standard; cDNA to mRNA; 990 BP.  |   |      |
| XX       | AAT04849;   |   |      |
| XX       | 28-MAY-1996 (first entry)   |   |      |
| XX       | Human adenosine 5' phosphosulphate kinase 990 bp gene fragment.                   |   |      |
| DE       | Human adenosine 5' phosphosulphate kinase; genetic disease; diagnosis; probe; ds. |   |      |
| KW       | Phosphosulphate kinase; genetic disease; diagnosis; probe; ds.                    |   |      |
| XX       | Homo sapiens.   |   |      |
| OS       | JP07250687-A.   |   |      |
| PN       | XX  |   |      |
| XX       | 03-OCT-1995.  |   |      |
| PD       | XX  |   |      |
| XX       | 14-MAR-1994; 94JP-0069861.  |   |      |
| PF       | XX  |   |      |
| XX       | 14-MAR-1994; 94JP-0069861.  |   |      |
| PR       | XX  |   |      |
| XX       | (NIRA ) UNITIKA LTD.  |   |      |
| PA       | XX  |   |      |
| XX       | WPI; 1995-370475/48.  |   |      |
| DR       | XX  |   |      |
| XX       | DNA fragment of human adenosine 5' phospho:sulphate kinase - used as              |   |      |
| PT       | a probe in the detection of genetic disease and for the production                |   |      |
| PT       | of 5' phospho:sulphate kinase   |   |      |
| XX       | Claim 1; Page 6; 6pp; Japanese.   |   |      |
| XX       | AAT04849 and AAT04850 are fragments of the human adenosine 5'                     |   |      |
| CC       | phosphosulphate kinase (5' PSK) of 990 and 600 bp respectively. "They             |   |      |
| CC       | may be used as probes for the detection of genetic disease. The 990               |   |      |
| CC       | bp fragment may also be used for low cost prodn. of human 5' PSK.                 |   |      |
| XX       | Sequence 990 BP; 281 A; 182 C; 267 G; 260 T; 0 other;                             |   |      |
| SQ       |   |   |      |

|    | Query Match           | 26.3%   | Score 484.8;        | DB 16;    | Length 990; |
|----|-----------------------|---|---------------------|-----------|-------------|
|    | Best Local Similarity | 68.8%;  | Pred. No. 6.2e-139; |           |             |
|    | Matches 680;          | Conservative 0;   | Mismatches 307;     | Indels 1; | Gaps 1;     |
| QY | 36                    | GCAGAAATCCACCAATGTAGTCTATCAGGCCCCACCATGTGCAGCAGGAATTAAGAGGGCA   | 95                  |           |             |
| Db | 3                     | GCAGAGAGCAACCAATGTCTACCTACCAAGCCCATCATGTGCAGGAGCAACAGAGAGGTCA   | 62                  |           |             |
| QY | 96                    | AGTGGTTTGGAAACAAAGGGGTGGGTTCCGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGG | 155                 |           |             |
| Db | 63                    | GGTGGTGGGACCAAGAGGTGGCTTTCTGTGTTGCACAGTTTGGCTTAACAGGCTTGTCTGG   | 122                 |           |             |

|    |     |   |      |
|----|-----|---|------|
| Qy | 156 | TCGTGGAAAAACAACGATAAGTTTTTGGCCCTGGAGGAGTACCTTGTCTCCCATGCCCATCC  | 215  |
| Db | 123 | AGCGGGAAGACTACTGTGAGCATGGCCCTGGAGGAGTACCTGTTGTTCATGGTATTC       | 182  |
| Qy | 216 | TTGTTACTCCTGGATGGGACAATGTCGTTCATGGCTTAACAGAAATCTCGGATTC         | 275  |
| Db | 183 | ATGCTACACTCTGGATGGTGACAATATTTCTGCTCAAGTCTCAATAAANAATCTTGGCTTTAG | 242  |
| Qy | 276 | TCCTGGGGACAGAGAGGAAAAATATCCGCCGGATTGTCTGAGTGGCTTAAGCTGTTTTGCTGA | 335  |
| Db | 243 | TCCTGAGACACAGAGAAGAAGATGTTCGACGGCATCGAAGATGTCTAAACTGTTTGCAGA    | 302  |
| Qy | 336 | TGCTGGTCTGCTCTGCATTACCAGCTTTATTTCTCCATTTCGCAAGGATCGTGAGATGC     | 395  |
| Db | 303 | TGCTGGCTTAGTGTGCATCAACAAGTTTCAATACCTTACACTCAGGATCGCAACAATGC     | 362  |
| Qy | 396 | CCGCAAAATACATGAATCAGCAGGGCTGCCATTCCTTTGAAAATATTTGTAGATGCACCTCT  | 455  |
| Db | 363 | AAGCAAAATTCATGAAGGTGCAAGTTTACCCTGTTTTTGAAGTATTTTGTGATGCTCCTCT   | 422  |
| Qy | 456 | AAATATTTGTCAAGACAGACAGCTAAAGAGCCTCTATAAAGGGCCAGCTGGGGAGAT       | 515  |
| Db | 423 | GCATGTTTGTCAACAGAGGAGTGTCAAGAGACTCTACAAAAGCCCGGCGAGGAGAAAT      | 482  |
| Qy | 516 | TAAAGGATTTACAGTATTGATCTTGATTTAGAAAGCTTGAAGCTCTGAGCTGTGCT        | 575  |
| Db | 483 | TAAAGTTTCACTGGATCGATTCTGAATATGAANAAGCAGAGCCCTGAGTTGGTCT         | 542  |
| Qy | 576 | TAAACCAATTTGTCCAGATGAGTGACTGTGTCCACAGGTAGTGGAACTTCTGCAGA        | 635  |
| Db | 543 | GAACACAGACTCCCTGTGTATTAATGACTGTGTGCAGCAAGTTGTGGAATCTTACAGA      | 602  |
| Qy | 636 | GCAGACATGTACCCATACTATAATCAAGATATCCAGCACTCTTTGTGCCGGAANA         | 695  |
| Db | 603 | ACGGGATTTGACTCTGGATGCACTTTATGAAGTAAAGAACATATATGTCCAGAAA         | 662  |
| Qy | 696 | CAAACTTGACACAGCTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCT      | 755  |
| Db | 663 | TAAACTTCATTTGGCAAAAACAGATCGGAAACATTTACCAGAGTGAAAATTAATAAGT      | 722  |
| Qy | 756 | GGATCTCAGTGGGTCCAGGTTTGGCGAAGGCTGGGCCACTCCCTCAAAGTTTCAAT        | 815  |
| Db | 723 | GGATGCAGTGGGTGCAGTTTTTGGAGAAGGTTGGGCAACCCCATTTGAATGGCTTAT       | 782  |
| Qy | 816 | GCGGGAAGAGTACTTACAGTTATGCATCTTTGACACCTGTCTAGATGATGGCGTAT        | 875  |
| Db | 783 | GAGAGAGGGAGTACTTGCAAGTCCCTTCAATTTTGATTGTTCTCTGGATGGAGGTGCTAT    | 842  |
| Qy | 876 | CAACATGAGCATCCCATTTGACTGCCCGTCTCTGCAGAGAGATAGACACGGCTGGAAG      | 935  |
| Db | 843 | TAACTTGCAGTACTATAGTTCGTGACTGCGACATCATGAAGATAAAGAGAGGCTGGATGG    | 902  |
| Qy | 936 | GTGCAGCAAGTTTTCTCTGGCACATGTTGGACGGAGGGTAGCTATCTTACGAGAGCGTGA    | 995  |
| Db | 903 | CTGTACAGCATTTGCTCTGATGTAT - GAGGCCCGCTGTGCCATCTCTCGCAATCCAGA    | 961  |
| Qy | 996 | ATTCTATGAACAGAAAAGAGGAACGC                                      | 1023 |
| Db | 962 | GTTTTTTGCACACAGGAAGAGGAGCG                                      | 989  |

|           |  |
|-----------|--|
| RESULT 10 |  |
| AAC34705  |  |
| ID        | AAC34705 standard; DNA; 1695 BP.                               |
| XX        | XX   |
| XX        | AAC34705;  |
| XX        | XX   |
| XX        | 17-OCT-2000 (first entry)                                      |
| XX        | XX   |
| DE        | Arabidopsis thaliana DNA fragment SEQ ID NO: 7588.             |
| XX        | XX   |
| XX        | Hybridisation assay; genetic mapping; gene expression control; |
| KW        | protein identification; signal transduction pathway;           |
| KW        |  |

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PF 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 06-MAY-1999; 99US-0132487.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.









```
99US-0145919.
99US-0145951.
99US-0146386.
99US-0146388.
99US-0146389.
99US-0147038.
99US-0147204.
99US-0147302.
99US-0147192.
99US-0147260.
99US-0147303.
99US-0147416.
99US-0147493.
99US-0147935.
99US-0148171.
99US-0148319.
99US-0148341.
99US-0148565.
99US-0148684.
99US-0149368.
99US-0149115.
99US-0149426.
99US-0149722.
99US-0149723.
99US-0149929.
99US-0149902.
99US-0149930.
99US-0150566.
99US-0150884.
99US-0151065.
99US-0151066.
99US-0151080.
99US-0151303.
99US-0151438.
99US-0151930.
99US-0152363.
99US-0153070.
99US-0153758.
99US-0154018.
99US-0154039.
99US-0154779.
99US-0155139.
99US-0155486.
99US-0155659.
99US-0156458.
99US-0156596.
99US-0157117.
99US-0157753.
99US-0157865.
99US-0158029.
99US-0158232.
99US-0158369.
99US-0159293.
99US-0159294.
99US-0159295.
99US-0159329.
99US-0159330.
99US-0159331.
99US-0159637.
99US-0159638.
99US-0159584.
99US-0160741.
99US-0160767.
99US-0160768.
99US-0160770.
99US-0160814.
99US-0160815.
99US-0160980.
99US-0160981.
99US-0160989.
99US-0161404.
99US-0161405.
99US-0161406.

27-JUL-1999;
28-JUL-1999;
02-AUG-1999;
02-AUG-1999;
02-AUG-1999;
03-AUG-1999;
04-AUG-1999;
04-AUG-1999;
05-AUG-1999;
05-AUG-1999;
06-AUG-1999;
06-AUG-1999;
09-AUG-1999;
09-AUG-1999;
10-AUG-1999;
11-AUG-1999;
12-AUG-1999;
13-AUG-1999;
13-AUG-1999;
16-AUG-1999;
17-AUG-1999;
18-AUG-1999;
20-AUG-1999;
20-AUG-1999;
20-AUG-1999;
23-AUG-1999;
23-AUG-1999;
25-AUG-1999;
26-AUG-1999;
27-AUG-1999;
27-AUG-1999;
27-AUG-1999;
30-AUG-1999;
31-AUG-1999;
01-SEP-1999;
07-SEP-1999;
10-SEP-1999;
13-SEP-1999;
15-SEP-1999;
16-SEP-1999;
20-SEP-1999;
22-SEP-1999;
23-SEP-1999;
24-SEP-1999;
28-SEP-1999;
29-SEP-1999;
04-OCT-1999;
05-OCT-1999;
06-OCT-1999;
07-OCT-1999;
08-OCT-1999;
12-OCT-1999;
13-OCT-1999;
13-OCT-1999;
14-OCT-1999;
14-OCT-1999;
14-OCT-1999;
14-OCT-1999;
18-OCT-1999;
21-OCT-1999;
21-OCT-1999;
21-OCT-1999;
22-OCT-1999;
22-OCT-1999;
23-OCT-1999;
25-OCT-1999;
25-OCT-1999;

99US-0161359.
99US-0161360.
99US-0161361.
99US-0161920.
99US-0161992.
99US-0161993.
99US-0162142.

Query Match          19.6%; Score 360.8; DB 21; Length 1687;
Best Local Similarity 58.6%; Pred. No. 1.9e-100;
Matches 686; Conservative 0; Mismatches 472; Indels 12; Gaps 3;

QY 676 GAACCTTTTGGCCGGAAACAAACTTGACCAGCTCCGAGCTGAGGCTGAAACTTCCT 735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 GAGCTCATCTGTGAAGACCAAGACGAGAGAAACACGACGCGCGGATTTGCCA 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 736 TCATTATCAATTACTAGCTCGATCTCCAGTGGGTCCAGGTTTTCAGCGAGGCTGGGCC 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 CGTGTGAGCTGACGGCGATTGACTTGCATTTGGATGTCATGTAAGCGAAGCTGGACA 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 796 ACTCCCTCAAAGGTTTCATCGGGGAGAAAGGACTTACAGGTTTATGACCTTTGACACC 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 AGTCCACTCGGAGGTTTATGAGAGAAATCCGAGTTCTCTCCAACTCTTCATTTAACTCG 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 856 CTGCTAGATGATG-----GGGTATCAACATGAGCATCCCACTTGTACTGCCGCTCTCT 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 CTAGCTTTGACGACGGCTCCGTCGTTAACTATGTCGCTGCTATTTGTTCTCGCTATTGAC 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 GCAGAGGATAGACACACGGCTGGAAGGTGACAGCAAGTTTGTCTGGGCACAT---GGTGA 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 GATGAACAAAAGCAGCTATCGGCGAGTCTACAGTGTGCTCTTTTCAATTCGATGGT 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 967 CGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGACGCTGT 1026
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 AACCCCTCTCTATCTCCAGCGATATTGAGATTTATAAGCATCCAAAGGAAGAGATA 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1027 TCCGCTGTTGGGGGACAAATGTACAAACACCCCATATCAAAATGGTGAT---GGAA 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 609 GCTAGAACATGGGGTACGACGGCTCCAGGTTTGCCTTACGTAGACGAGCGGATAACTAAT 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1084 AGTGGGACTGGCTGGTGGTGAGACCTTCAGGTGCTGGAGAAAATAAGATGAATGAT 1143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 669 GCTGGAACCTGGCTCATTTGGGGGTGATCTTGAGGTTCTTGAGCCAGTGAAGTACAAATGAT 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1144 GGGCTGGAACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAAGAAATGAAT 1203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 729 GGGCTTGATCGTTTCAGGCTTTCGCTGCTGAGTTACGTAAGAGTTGGAGAAGCGTAAT 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1204 GCTGATCGGCTGTTTGATTTCCAGTTCGCGCAATCCTGTCCACATGCCCATGCCCTGTG 1263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 789 GCGGATGCGGTGTTTGCTTTTCCAGCTGAGGAATCCTGTTCATAATGTGCTCTTCTTT 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1264 ATGAGGACACCTGCCGACGGCTCTAGAGAGGGGCTACAGCAACCGGCTCCTCTACTA 1323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 849 ATGACTGATACTCGTAGGAGACTTCTTGAGATGGGTTACAAAACCTATCTTTCTTCCTT 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1324 CACCTCTTGGCGGCTGACCAAGATGACGATGTGCTCTCTAGACTGCGCGATGAAGCAG 1383
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 909 CATCCGTTAGGTGGGTTTACAAAAGCTGATGTCTTCTTTAGATTGGAGGATGAAGCAA 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1384 CACGCGCTGTGCTCGAGGAAGGGTCTCGATCCCAAGTCAACCACTTTGTTGCCATCTTT 1443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 969 CACGAGAAGGTTCTAGAGGATGGTGTCTCGATCCGAGACTCAGTGGTGTTCGATATTC 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1444 CCGTCTCCCATGTTATGCTGGCCCAACAGAGTCCAGTGGCAGTCCAGGTCGCCGATG 1503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1029 CCGTCACTATGATACGCTGGTCCAAACCAAGTGCAGTGGCAGCAAGAGGTAGATC 1088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1504 ATTGCGGTCGCAATTTCTATATTTGGGAGGAGACCCCTCAGGAATGCCCATCTCTGAA 1563
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1089 AATGCTGGTCTAACTTTTACATTTGTTGGTGTGATCTCTGCTGGATGGGTCTATCCAGTA 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1564 ACCAAGAAGGATCTGTATGNAACCACTCATGCGGGGCAAGGCTCTTGAGCATGCGCCCTGGC 1623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
Db 1149 GAGAAACGTGATCTTTACGATCTCTGATCATGGAAGAAAGTACTAAGCATGCGACCAGGA 1208
QY 1624 CTCACCTCTGTGGAATCATTCATTCCGAGTGCTGCTCAACAAGCCAAAAGGCC 1683
Db 1209 CTCGACGACCTCAACATCTCTCTTACGGTGTGCTGATATGACAAGCAGCGAAGCAAG 1268
QY 1684 ATFGACTTCTATGATCCAGCAAGGACAAATGAGTTTGACTTCATCTCAGGAACCTGGAATG 1743
Db 1269 ATGGCTTCTTCGATCCCTCGAGGCTCAAGATTCTTGTTCATCTCCGGCACTAAGATG 1328
QY 1744 AGGAAGCTCGCCCGGAGGAGAGAAATCCCCAGATGGCTTCATGCCCCCAAGCAATGG 1803
Db 1329 CCACATTTGGCAAGAAACAAACGAAACCCGACAGCGTTTATGTGCCCAAGGTGGATGG 1388
QY 1804 AAGGTCTCTGACAGATATTATACAGTCCCTG 1833
Db 1389 AAGGTCTCTGATGATTACTATGAGAGCTTG 1418

RESULT 15
AAC42592
ID AAC42592 standard; DNA; 1533 BP.
XX
AC AAC42592;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36133.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
```

```
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.5%; Score 359.8; DB 21; Length 1533;
Best Local Similarity 58.7%; Pred. No. 3.7e-100; Indels 12; Gaps 3;
Matches 682; Conservative 0; Mismatches 467;

Qy 676 GAACCTCTTTGTCGCGGAAACAAACTTGCACCACGTCGCCAGCTGAGGCTGAAACTCTCCCT 735
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 GATCTTGTGTACCGAACCGAGACGGCGAGAGAAGAAACACGACGCGCGGATTTGCCG 378

Qy 736 TCATTATCAATTACTAAGCTGGATCTCCAGTGGGTCCAGGTTTTCAGGGAAGCTGGGCC 795
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 AGAGTGAGATTGACGCGCATTTGATTTGCAATGGATGCTGTTGTTGAGTGGAGTTGGCT 438

Qy 796 ACTCCCTCTAAAGGTTTCATCGCGGAGAGGAGTACTACAGGTTATGACCTTTGA---- 851
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 AGTCCTCTTCGTGGGTTTATGAGGGAATCTGAGTTCCTCCAACTCTTCATTCAATTTG 498

Qy 852 --CACCTCTAGATGATGGCGTATCAACATAGCATCCCATTTGACTGCCGCTCTCT 909
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 TTGAATCTCGATGATGGTCTGTTGTTAAATATGCTGTGCTTATTTCTTCGCGATTGAT 558

Qy 910 GCAGAGATAGACACGCGCTGGAAGGTGCAGCAAGTTTGTCTGGGCACATGGTGGACGG 969
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GATCAACAAAAGCCCTAATCGGTGAATCTAAACGTGTCTCCCTTGTGTTCTGTATGAT 618

Qy 970 AG---GGTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGAACGCTGT 1026
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 AATCCAATCGTATTTCAATGATATTGAGATTTTAAACATCCGAAGAAGAGCGAATA 678

Qy 1027 TCCCGTGTGGGGGACAAACATGTACAAAACACCCCATATCAAAATGGTGAT---GGAA 1083
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCGAGAACTTGGGCTACGACTGCACCGGTTTGGCTTATGTAGAGAGGCGATAACCAAT 738

Qy 1084 AGTGGGACTGGCTGGTGGTGGAGACCTTCAGTGTCTGGAGAAATTAAGATGAATGAT 1143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 GCTGGAGACTGGCTCATTTGGGGGTGATCTTGGAACTTTTGAACCTGTTAAGTACAATGAT 798

Qy 1144 GGGCTGGACCAATACCCGCTGCACACCTCTGAGCTCAACAGAAATGTAAGAAATGAAT 1203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 GGGCTTGATCGTTTCAGGCTTTCCCGCTTTGAACTGCGTAAAGAGCTAGAGAAACGTGGT 858

Qy 1204 GCTGATCGGTGTTTGCAATCCAGTTGCGCAATCTGTCCACAATGCCATGGCCTGTG 1263
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 GCGGATCGGTCTTTGGCTTTACAGTTAGGAAACCCAGTTTCATATATGACATGCTCTCTT 918

Qy 1264 ATGACGACACCTGCCCGCAGCTCCTAGAGAGGGCTACAGCACCGGCTCCTCTACTA 1323
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 ATGACTGATATCTGATAGGAGACTTCTTGATGATGGTTTATAAAACCTATCTCTTTGCTT 978

Qy 1324 CAGCCTCTGGCGGCTGGACCAAGGATGACGATGTGCTCTAGACTGGCGGATGAAGCAG 1383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 CATCCATTTGGAGGGTTTACAAAGCTGATGATGTTCTCTAGCTGGCGAATGAACAG 1038

Qy 1384 CACGGGCTGTCTCGAGGAAGGGTCTCGATCCCAAGTCAACCATTTGTCATCTTT 1443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 CAGGAGAAAGGTGTAGAGGATGGTGTCTTGTATCCAGAGACTACTGTGGTTTTCATATTC 1098

Qy 1444 CCGTCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGGACTCGAGTCCCGGATG 1503
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1099 CCATCTCCAAATGCTATGCTGGTCCCAACCGAGTCCATGAGCCGAAAGGCTAGGATG 1158

Qy 1504 ATTTGCGGGTGCCAAATTTTCTACATTTGGGGAGGAGACCTTCAGGAATGCCCATCTGAA 1563
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1159 AATGCTGGTCTAACTTCTACATTTGTCGGTAGGATCCCGGCTGGAATGGGTGCTCCGCTG 1218

Qy 1564 ACCAAGAAGGATCTGTATGAACCCATCATATGGGGGCAAGTCTTTGAGCATGGCCCTGGC 1623
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1219 GAGAAACGATGCTGTACGATGTGATACGCGGGAAGAAGTTCTTAAGCATGGTCTCTGGA 1278

Qy 1624 CTCACCTCTCTGGAAATCATTTCCATTCGAGTGGCTGCCTACAACAAGCAAAAGGCC 1683
```

```

Db 1279 CTCGAACGACTCAACATTCTTCCTTTCAGGGTGTGCTACGATAAGACACAAGGAAAG 1338
QY 1684 ATGGACTTCTATGATCCAGCAAGGCAATGAGTTTGACTTCATCTCAGGAACTCGAATG 1743
Db 1339 ATGGCTTTCTTTGATCCCTCAAGGGCTCAGGACTTCTTGTTCATTTCTGGCACAAGATG 1398
QY 1744 AGAAGCTCGCCCGGAAGGAGAGATCCCCAGATGGCTTCATGGCCCCCAAGCATGG 1803
Db 1399 AGAGCATTTGGCAAGAAACAGAGAACCCACCAGATGGATTATGTGCCCTGGAGGCTGG 1458
QY 1804 AAGGTCCTCACAGATTATTAC 1824
Db 1459 AAGTCCCTGTGGATTACTAC 1479

```

Search completed: June 4, 2003, 15:23:01  
Job time : 457 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 15:13:56 ; Search time 94 Seconds  
(without alignments)  
6019.347 Million cell updates/sec

Title: US-09-898-165B-9

Perfect score: 1845

Sequence: 1 atgtcggggatcaagaagca.....ggctcctggagaagaactaa 1845

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description         |
|------------|-------|-------------|---------|----|---------------------|
| 1          | 990   | 53.7        | 2506    | 1  | US-08-879-561-4     |
| 2          | 383   | 20.8        | 1697    | 4  | US-09-346-408-5     |
| 3          | 368.2 | 20.0        | 1890    | 4  | US-09-346-408-7     |
| 4          | 368   | 19.9        | 2280    | 4  | US-09-346-408-3     |
| 5          | 226   | 12.2        | 2323    | 4  | US-09-149-476-24    |
| 6          | 197.4 | 10.7        | 516     | 4  | US-09-346-408-1     |
| 7          | 169.2 | 9.2         | 1160    | 4  | US-09-153-310-41    |
| 8          | 143   | 7.8         | 563     | 4  | US-09-346-408-9     |
| 9          | 123   | 6.7         | 615     | 4  | US-09-134-001C-1680 |
| 10         | 110   | 6.0         | 4411529 | 4  | US-09-103-840A-1    |
| 11         | 108.4 | 5.9         | 4403765 | 4  | US-09-103-840A-2    |
| 12         | 46.6  | 2.5         | 7218    | 1  | US-08-232-463-14    |
| 13         | 34.2  | 1.9         | 289     | 4  | US-09-007-005-17    |
| 14         | 34.2  | 1.9         | 289     | 4  | US-09-244-796-17    |
| 15         | 34.2  | 1.9         | 3942    | 4  | US-09-162-484-19    |
| 16         | 33.8  | 1.8         | 4403765 | 4  | US-09-103-840A-2    |
| 17         | 33.8  | 1.8         | 4411529 | 4  | US-09-103-840A-1    |
| 18         | 33.6  | 1.8         | 5474    | 4  | US-09-040-738-1     |
| 19         | 33.6  | 1.8         | 5474    | 4  | US-08-652-426A-1    |
| 20         | 33.6  | 1.8         | 5943    | 1  | US-08-206-176-1     |
| 21         | 33.6  | 1.8         | 10411   | 4  | US-08-961-527-89    |
| 22         | 33.2  | 1.8         | 3807    | 1  | US-08-357-598-1     |
| 23         | 33.2  | 1.8         | 3807    | 1  | US-09-003-289-1     |
| 24         | 33.2  | 1.8         | 3807    | 5  | PCT-US95-16435-1    |
| 25         | 32.8  | 1.8         | 3726    | 1  | US-08-173-497-1     |
| 26         | 32.8  | 1.8         | 3726    | 1  | US-08-286-889-1     |
| 27         | 32.8  | 1.8         | 3726    | 1  | US-08-485-618-1     |

|    |      |     |      |   |                   |
|----|------|-----|------|---|-------------------|
| 28 | 32.8 | 1.8 | 3726 | 1 | US-08-362-652-1   |
| 29 | 32.8 | 1.8 | 3726 | 1 | US-08-605-672-1   |
| 30 | 32.8 | 1.8 | 3726 | 2 | US-08-482-293A-1  |
| 31 | 32.8 | 1.8 | 3726 | 2 | US-08-943-363-1   |
| 32 | 32.8 | 1.8 | 3726 | 4 | US-09-193-043-1   |
| 33 | 32.8 | 1.8 | 3726 | 4 | US-09-688-307A-1  |
| 34 | 32.8 | 1.8 | 3785 | 1 | US-08-485-618-98  |
| 35 | 32.8 | 1.8 | 3785 | 1 | US-08-605-672-98  |
| 36 | 32.8 | 1.8 | 3785 | 2 | US-08-482-293A-98 |
| 37 | 32.8 | 1.8 | 3785 | 2 | US-08-943-363-98  |
| 38 | 32.8 | 1.8 | 3785 | 4 | US-09-193-043-98  |
| 39 | 32.8 | 1.8 | 3785 | 4 | US-09-688-307A-98 |
| 40 | 32.8 | 1.8 | 3956 | 1 | US-08-485-618-97  |
| 41 | 32.8 | 1.8 | 3956 | 1 | US-08-605-672-97  |
| 42 | 32.8 | 1.8 | 3956 | 2 | US-08-482-293A-97 |
| 43 | 32.8 | 1.8 | 3956 | 2 | US-08-943-363-97  |
| 44 | 32.8 | 1.8 | 3956 | 4 | US-09-193-043-97  |
| 45 | 32.8 | 1.8 | 3956 | 4 | US-09-688-307A-97 |

#### ALIGNMENTS

RESULT 1  
US-08-879-561-4  
; Sequence 4, Application US/08879561  
; Patent No. 5817482  
; GENERAL INFORMATION:  
; APPLICANT: Bardman, Olga  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Hawkins, Phillip R.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/879,561  
; FILING DATE: Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0325 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2506 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: LUNGN02  
; CLONE: 373887  
US-08-879-561-4









Db 1262 GATAAACTCAGGCAAAATGGCCCTTTTCGATCCCTCGAGACCTCAGGATTTCTCTTTTC 1321  
QY 1726 ATCTCAGCAACTCGAATCAGGAAGCTCGCCGGGAAGGAGAGATCCCCAGATGGCTTC 1785  
Db 1322 ATATCCGCGCACCAGATGAGACACTACCAAGAACAAAGAGATCCACCAGAGGGTTT 1381  
QY 1786 ATGGCCCCCAAGCATGGAAGTCTCTGACAGATTATTA 1823  
Db 1382 ATGTGCCCCAGGGATGCAAGGTGTGTGTTGAATATTA 1419

## RESULT 5

US-09-149-476-24  
; Sequence 24, Application US/09149476  
; Patent No. 6420526  
; GENERAL INFORMATION:  
; APPLICANT: Rosep et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002PI  
; CURRENT APPLICATION NUMBER: US/09/149,476  
; CURRENT FILING DATE: 1998-09-08  
; EARLIER APPLICATION NUMBER: PCT/US98/04493  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 60/040,162  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,333  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/038,621  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,626  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,334  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,336  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/040,163  
; EARLIER FILING DATE: 1997-03-07  
; EARLIER APPLICATION NUMBER: 60/047,600  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,615  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,597  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,502  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,633  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,583  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,617  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,618  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,503  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,592  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,581  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,584  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,500  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,587  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,492  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,598  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,613  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,582  
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,612  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,632  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/047,601  
; EARLIER FILING DATE: 1997-05-23  
; EARLIER APPLICATION NUMBER: 60/043,580  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,568  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,314  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,569  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,311  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,671  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,674  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,669  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,312  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,313  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,672  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/043,315  
; EARLIER FILING DATE: 1997-04-11  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/056,886  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,877  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,889  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,893  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,630  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,878  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,662  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,872  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,882  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,637  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,903  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,888  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,879  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,880  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,894  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,911  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,636  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,874  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,910  
; EARLIER FILING DATE: 1997-08-22  
; EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 12.2%; Score 226; DB 4; Length 2323;  
Best Local Similarity 92.2%; Pred. No. 1.4e-61;  
Matches 271; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 1 ATGTCGGGATCAAGAAGCAACGAGGAGACGAGAAATCCACCAATGTAGTCTAT 60

Db 165 ATGTCGGGATCAAGAAGCAACGAGGAGACGAGAAATCCACCAATGTAGTCTAT 224  
QY 61 CAGGCCACCATGTGACGAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGTTC 120  
Db 225 CAGGCCACCATGTGACGAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGTTC 284  
QY 121 CGAGGATGTACCGTGTGGCTTAACAGGCTCTCTCTGGTGTCT -GGAAAAACAACGATAAGTTT 179  
Db 285 CGAGGATGTACCGTGTGGCTTAACAGGCTCTCTCTGGTGTGGAAAAACAACGATAAGTTT 344  
QY 180 TGCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTGTGTACTCTCCCTGATGGGACAA 239  
Db 345 TGCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCT -TGTTAATTCCTGATGGGACAA 402  
QY 240 TGTCGGTGTGGCTTAACAGAAATCTCGGATCTCTCTCTGGGACAGAGGA 293  
Db 403 TGTCGGTGTGGCTTAACAGAAATCTCGGATCTCTCTCTGGGACAGGA 456

## RESULT 6

US-09-346-408-1  
; Sequence 1, Application US/09346408B  
; Patent No. 638966  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Anderson, Shawn  
; APPLICANT: Falco, Carl  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
; FILE REFERENCE: BB-1167-A  
; CURRENT APPLICATION NUMBER: US/09/346,408B  
; CURRENT FILING DATE: 1999-07-01  
; EARLIER APPLICATION NUMBER: 60/092,833  
; EARLIER FILING DATE: July 14, 1998  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 516  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (479)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (495)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (508)  
US-09-346-408-1

Query Match 10.7%; Score 197.4; DB 4; Length 516;  
Best Local Similarity 65.6%; Pred. No. 7e-53;  
Matches 320; Conservative 0; Mismatches 162; Indels 6; Gaps 2;

QY 1140 TGATGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAATGTAAGAAAT 1199  
Db 1 TGATGGCTAGATCAGTATCGCTCTCCAGCACAGCTCGGTGAAGAGTTTGCCAGGG 60  
QY 1200 GAATGCTGATGCGGTGTTTGCATTCAGTTGCGCAATCTTCACAAATGGCCATGCCCT 1259  
Db 61 CAATGCTGATGCAATTTTGCCTTCAGCTTCGCAATCCAGTACACAATGGCATGCTCT 120  
QY 1260 GTTGATGACGAGCACCTCGCGCAGGCTCTAGAGAGGGGTACAAAGCACCCTGCTCTCT 1319  
Db 121 ACTTATGACGAGCACAGGCAACGCTCTCTTGAGATGGGTATATAAACCTGTTCTCT 180  
QY 1320 ACTACACCTCTGGCGGCTGGACCAAGGATGACGATGCTCTAGACTGGCGGATGAA 1379  
Db 181 GCTCCATCCACTGGGAGGATTACAAAGAGCAGATGATGCTCTCTAGTTGAGAGATGAA 240  
QY 1380 GCAGCAGCGGCTGCTCTCGAGGAAGGGTCTTGGATGCCCAAGTCACCAATGTTGCCAT 1439

Db 241 GCAACATGAGAGGTTCTTGAGAGAGGTGCTCAACCCAGAAATCAACTGTTGCGAT 300  
 QY 1440 CTTTCGGTCCCGATGATATGCTGGCCGCCACAGAGTCCAGTGGCAGTGCAGGTCGCC 1499  
 Db 301 CTTCCCTCCTCATGATGCTGGCCCACTGAGTTCAATGCGATCCTAAGGCTCG 360  
 QY 1500 GATGATGGGGTGCCAAATTCATATGTTGGGAGGAGCCCTGCAGGAATGCCCATCC 1559  
 Db 361 TATTAATGCTGGCGCAAAATTCATATGTTGGAAGGATCTGCTGTATGAGCACCCAC 420  
 QY 1560 TGAACCAAGAAGATCTGATGAACCCACTCATGCGGGGCAAGTCTTGAGCATGGCCC 1619  
 Db 421 AGAGA-----AAGAGGACTCATGCTGATCATCGGAAGAAGGT-TTGAGCATGGCTCC 474  
 QY 1620 TGGCCTCA 1627  
 Db 475 TGCNTCA 482

RESULT 7

US-09-153-310-41  
 ; Sequence 41, Application US/09153310  
 ; Patent No. 6326184  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gjermansen, Claes  
 ; Hansen, Jorgen  
 ; Johannesen, Pia Francke  
 ; Pedersen, Mogens Bohl  
 ; Sorensen, Steen Bech  
 ; TITLE OF INVENTION: Method of producing a composite  
 ; fermented beverage using genetically modified yeast  
 ; strains  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/153,310  
 ; FILING DATE: 15-Sep-1998  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: <Unknown>  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Bent, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: <Unknown>  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-672-5300  
 ; TELEFAX: 202-672-5399  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 41:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1160 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Genomic DNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
 US-09-153-310-41

Query Match 9.2%; Score 169.2; DB 4; Length 1160;  
 Best Local Similarity 57.6%; Pred. No. 1.1e-43;

Matches 303; Conservative 0; Mismatches 223; Indels 0; Gaps 0;  
 QY 124 GGATGACCGTGGGTAAACAGGTCTCTCTGGTCTGAAAAACAAGTAAAGTTTGGCC 183  
 Db 447 GGCCTGACCGTTGGTTGACCGGTCTAAGTGGCTCAGAAAAAGTACAATAGCTTGCA 506  
 QY 184 CTGAGGAGTACCTTGTCTCCCATGCGATCCCTTGTGTCTCCCTGGAGGACAAATGTC 243  
 Db 507 CTGGAACAATTAAGTCTCAAAAAAATTAAGTCTTATAGTTAGTGTGATAACATT 566  
 QY 244 CGTCATGGCCCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAAAATATCCG 303  
 Db 567 CGTTTGGTTGAATAAGGATTTGGGCTTCTCAGAAAAAGACAGAAATGAAAAACATTCGT 626  
 QY 304 CGGATTCGCTGAGGTGGCTAAAGCTGTTTCTGCTGATGCTGCTGCTGCTGCTGCTGCT 363  
 Db 627 AGAATTAGTCAAGTATCAAGCTATTCGCTGATTCGCTGATTCGCTGATTCATTCATT 686  
 QY 364 ATTTCTCCATTCGCAAGGATCGTGAGAAATGCGCGGCAAAATACATGAATCAGCAGGCTG 423  
 Db 687 ATTTCCCATACAGAGTCGATAGACAGAGAGCCGCTGATTTACATAAGGAGCAGGCTTG 746  
 QY 424 CCATTCTTTGAAATATTTGATAGTGCACCTCTAAATATTTGTGAACGACAGAGCTAAAA 483  
 Db 747 AGTTTCATTCGAAATTTTGTGATGTTTCAATTAAGTCTGCTGAGCAAGAGACCCCTAAG 806  
 QY 484 GGCCTCTATAAAAGGGCCAGAGCTGGGAGATTAAGGATTTACAGGTATTTGATTCGTAT 543  
 Db 807 GGTTGTATTAAGAACCCAGAGAGGTGCTATTAAGAGTTCAGTGTATTTTCAGCTCCT 866  
 QY 544 TATGAGAACTGAAACTCTCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAGTGAC 603  
 Db 867 TAGCAAGCTCCAAAGGCCAGAGTTGCAATTTAAGAACTGACCAAAAGACTGTTTGAAGAA 926  
 QY 604 TGTGTCACCAAGTAGTGGAACTTCTCAGAGCAGACATTCGTAC 649  
 Db 927 TGTGCTGCTATCATTTATGAGTACCTGCTCAATGAGAAAGATTATCC 972

RESULT 8

US-09-346-408-9  
 ; Sequence 9, Application US/09346408B  
 ; Patent No. 6338966  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Allen, Steve  
 ; APPLICANT: Anderson, Shawn  
 ; APPLICANT: Falco, Carl  
 ; APPLICANT: Rafalski, Antoni  
 ; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins  
 ; FILE REFERENCE: BB-1167-A  
 ; CURRENT APPLICATION NUMBER: US/09/346,408B  
 ; EARLIER FILING DATE: 1999-07-01  
 ; EARLIER APPLICATION NUMBER: 60/092,833  
 ; EARLIER FILING DATE: July 14, 1998  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 9  
 ; LENGTH: 563  
 ; TYPE: DNA  
 ; ORGANISM: Triticum aestivum  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (337)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (404)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (437)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (475)  
 ; FEATURE:









**THIS PAGE BLANK (user)**



| Query Match           | 98.3%;          | Score 1813.6;  | DB 9;      | Length 2617; |
|-----------------------|-----------------|--|------------|--------------|
| Best Local Similarity | 99.0%;          | Pred. No. 0;   |            |              |
| Matches 1841;         | Conservative 0; | Mismatches 4;  | Indels 15; | Gaps 1;      |
| QY                    | 1               | ATGTCGGGGATCAAGAAAGCAAAAGACGGAGAACCCAGCAGAGAAATCCACCAATGTAGTCTAT | 60         |              |
| DB                    | 93              | ATGTCGGGGATCAAGAAAGCAAAAGACGGAGAACCCAGCAGAGAAATCCACCAATGTAGTCTAT | 152        |              |
| QY                    | 61              | CAGGCCCCACCATGTGACGAGGAATTAAGAGAGGCCAAGTGGTTGGAACAAGGGGTGGGTTTC  | 120        |              |
| DB                    | 153             | CAGGCCCCACCATGTGACGAGGAATTAAGAGAGGCCAAGTGGTTGGAACAAGGGGTGGGTTTC  | 212        |              |
| QY                    | 121             | CAGGATGTACCGTGTGGCTTACACAGGTCTCTCTGGTCTCGTGAACCAACGATAAGTTTT     | 180        |              |
| DB                    | 213             | CAGGATGTACCGTGTGGCTTACACAGGTCTCTCTGGTCTCGTGAACCAACGATAAGTTTT     | 272        |              |
| QY                    | 181             | GCCTGTGAGGAGTACCTTGTCTCCCATGCOATCCCTTGTCTACTCCCTGGATGGGGACAAT    | 240        |              |
| DB                    | 273             | GCCTGTGAGGAGTACCTTGTCTCCCATGCOATCCCTTGTCTACTCCCTGGATGGGGACAAT    | 332        |              |

QY 241 GTCGTCATGGCTTTACAGAAATCTCGAATTCCTCTCTGGGACAGAGAGAAATATC 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
333 GTCGTCATGGCTTTACAGAAATCTCGAATTCCTCTCTGGGACAGAGAGAAATATC 392  
QY 301 GCGCGGATTTGCTGAGGTGGCTAAAGCTGTTCTGATGCTGCTGCTGCTATACCAGC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
393 GCGCGGATTTGCTGAGGTGGCTAAAGCTGTTCTGATGCTGCTGCTGCTATACCAGC 452  
QY 361 TTTATTTCTCCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATCAAGCAGG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
453 TTTATTTCTCCATTCGCAAGGATCGTGAGATGCCCGCAAAATACATCAAGCAGG 512  
QY 421 CTGCCATTTCTTGAATATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGAGACGTA 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
513 CTGCCATTTCTTGAATATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGAGACGTA 572  
QY 481 AAAGCCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGTTATGATCT 540  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
573 AAAGCCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGTTATGATCT 632  
QY 541 GATTATGAGAACTTGAACCTCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAGT 600  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
633 GATTATGAGAACTTGAACCTCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAGT 692  
QY 601 GACTGTGTCCACAGGTAGTGGAACTTCTGCAAGAGCAGAAATTTGTACCCCTATACATA 660  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
693 GACTGTGTCCACAGGTAGTGGAACTTCTGCAAGAGCAGAAATTTGTACCCCTATACATA 752  
QY 661 ATCAAGATATCCAGCACTCTTTGTCCGGAACCAATTTGACCAGCTCCGAGCTGAG 720  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
753 ATCAAGATATCCAGCACTCTTTGTCCGGAACCAATTTGACCAGCTCCGAGCTGAG 812  
QY 721 GCTGAACTCTCCCTTCATTAATTAATTAAGCTGATCTCCAGTGGGTCCAGGTTTG 780  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
813 GCTGAACTCTCCCTTCATTAATTAATTAAGCTGATCTCCAGTGGGTCCAGGTTTG 872  
QY 781 AGCGAAGCTGGGCACTCCCTCTCAAGGTTTCATGCGGGAGAGAGTACTTACAGTT 840  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
873 AGCGAAGCTGGGCACTCCCTCTCAAGGTTTCATGCGGGAGAGAGTACTTACAGTT 932  
QY 841 ATGCACCTTTGACACCTCTGCTAGA-----TGATGGCTGTACCAATGAGC 885  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
933 ATGCACCTTTGACACCTCTGCTAGATGGCATGGCCCTTCTGATGGCTGTACCAATGAGC 992  
QY 886 ATCCCATTTGACTGCCCGTCTCTGCAGAGGATAAGACACGGCTGGAAGGTGCAGCAAG 945  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
993 ATCCCATTTGACTGCCCGTCTCTGCAGAGGATAAGACACGGCTGGAAGGTGCAGCAAG 1052  
QY 946 TTTGCTCTGACATGTGGACGAGGAGTAGCTATCTACGAGACGCTGAATTCATGAA 1005  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1053 TTTGCTCTGACATGTGGACGAGGAGTAGCTATCTTACGAGACGCTGAATTCATGAA 1112  
QY 1006 CACAGAAAAGGAAAGCTGTTCCCGTGTGTTGGGGCAACATCTACAAACACCCCAT 1065  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1113 CACAGAAAAGGAAAGCTGTTCCCGTGTGTTGGGGCAACATCTACAAACACCCCAT 1172  
QY 1066 ATCAAAATGTGATGGAAGTGGGACTGCTGTTGGTGAGACCTTCAGGTGCTGGAG 1125  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1173 ATCAAAATGTGATGGAAGTGGGACTGCTGTTGGTGAGACCTTCAGGTGCTGGAG 1232  
QY 1126 AAAATAGATGATGAGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAG 1185  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1233 AAAATAGATGATGAGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAG 1292  
QY 1186 AAATGTAAGAAATGAATGCTGATGCGGTGTTGCAATTCAGTTGCGCAATCTCTGTCAC 1245  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1293 AAATGTAAGAAATGAATGCTGATGCGGTGTTGCAATTCAGTTGCGCAATCTCTGTCAC 1352  
QY 1246 AATGGCATGCCCTGTTGATGCAGGACACCTGCCGAGGCTCTCTAGAGAGGGGTACAAG 1305  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1353 AATGGCATGCCCTGTTGATGCAGGACACTCGCGGAGGCTCTCTAGAGAGGGGTACAAG 1412

## RESULT 2

US-09-984-245-86  
; Sequence 86, Application us/09984245  
; Patent No. US20020165374A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004P1  
; CURRENT APPLICATION NUMBER: US/09/984,245  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30



QY 1629 CTCTGCTGAATCATTCCTCCGAGTGGCTTACACAAAGCCAAAGCCATCGA 1688  
Db 1714 CACTTTGGAATAGTTCCTTTTCAGTTGCAGTTTACACAAAGAAAGAGCGTATGA 1773  
QY 1689 CTTCTATGATCAGCAAGGCAATAGTTTGACTTTCATCTCAGGAATCGAATGAGAA 1748  
Db 1774 CTACTATGACTGTAACACCATGAGACTTTGAATTTATTCAGGAACAGATCGGAA 1833  
QY 1749 GCTCGCCCGGAAGAGAGATCCCCAGATGGCTTCATGCCCCCAAGAGTGAAGT 1808  
Db 1834 ACTTGCTCGAAGGCCAGAACACCTGGAAGTTTCATGGCTCCCAAGGCTTGACCGT 1893  
QY 1809 CCTGACAGATTATTACAGTCCCTGGAGAA 1838  
Db 1894 GCTGACAGAACTACTACAAATCCTTGGAGAA 1923

## RESULT 3

US-09-984-245-116  
; Sequence 116, Application US/09984245  
; Patent No. US20020165374A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: PZ004P1  
; CURRENT APPLICATION NUMBER: US/09/984,245  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,154  
; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: US 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: US 60/060,862  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 343  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 116  
; LENGTH: 2527  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-984-245-116

Query Match 53.7%; Score 990; DB 9; Length 2527;  
Best Local Similarity 71.3%; Pred. No. 0;  
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;  
QY 9 GATCAAGAAGCAAAAGACGGAGAACACGACAGAAATCCACCAATGTAGTCTATCAGGCCCA 68  
Db 94 GAGCAATAACCGCGCAGAACTGGGAAATGCAGAGAGCAACCAATGTACCTACCAAGCCCA 153  
QY 69 CCATGTGAGCAGGAATAAGACAGGCGCAAGTGTGGAAACAAGGGTGGTCCGAGGATG 128  
Db 154 TCATGTCAGCAGGAAACAAGAGAGGTGAGTGGTGGGACAGAGGTGGCTTTCTGGTGTG 213  
QY 129 TACCGTGTGGCTAACAGAGTCTCTCTGTGTGGTGGAAACAAACAGATAAGTTTTGCCCTGGA 188  
Db 214 CACAGTTTGGCTAACAGAGTGTCTGGAGCGGAAAGACTACTGTGACATGGCTTGA 273  
QY 189 GGAGTACTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGGTCA 248  
Db 274 GGAGTACTGTGTGTATTCATGTATTCATCTGATGTGATGATGATGATGATGATGATG 333  
QY 249 TGGCTTTAAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATATCCCGCGAT 308  
Db 334 AGGTCTCAATAAATCTTGGCTTTAGTCTTGAAGACAGAGAGAAATGTTCCAGCGAT 393  
QY 309 TGCTGAGGTGGCTAAGCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 368  
Db 394 CGCAGAAAGTTGCTAACTGTTTGCAGATGCTGGCTTAGTGTGATCATCAAGTTTCATATC 453  
QY 369 TCCATTTCGCAAGGATCGTGAGATGCCCAAAATACATCAATCAGCAGGCGTCCCAT 428  
Db 454 ACCTTACACTCAGGATCGCAACAATCAAGGCAAAATTCATGAAGGTGCAAGTTTACCGTT 513  
QY 429 CTTTGAATAATTTGTAGATGCACCTCTAAATATTTGTGAAGACAGACAGCTTAAAGGCGCT 488  
Db 514 TTTTGAAGTATTTGTTGATGCT 573  
QY 489 CTATAAAGGCGCAGAGCTGGGAGATTAAGAGATTTACAGGTATTTGATTTCTGATTTGA 548  
Db 574 CTACAAAAAGCCCGCAGAGAAATTAAGGTTTCACTGGGATCGATTTCTGAATATGA 633  
QY 549 GAAACCTGAAACTCTCGAGCTGTGCTTAAACCAATTTGCCACACTGAGTACTGTGT 608  
Db 634 AAAGCCAGAGCGCCCTGAGTTGGTGTGAAGACAGACTCTCTGTGATGTAATGATGCTGT 693  
QY 609 CCACCAAGTGTGGAATCTTCTGAAGACAGAAATTTGTACCTTATCTACTATAATCAAGA 668  
Db 694 CCAGCAAGTTTGGAACTTCTACAGGAACGGGATATTTGTACCTGTGGATGCATCTTATGA 753  
QY 669 TATCCAGCAACTCTTTGTGCGGAAACAACATTTGACCACTGCGAGCTGAGGCTGAAC 728  
Db 754 AGTAAAGAACTATATGTGCCAGAAATAAATTCATTTGGCAAAAAACAGATGCGGAAC 813  
QY 729 TCTCCCTTCAATTAATTAAGTGTGATCTCCAGTGGTCCAGGTTTGTAGCGAAGG 788  
Db 814 ATTACAGCACTGAAATTAATTAAGTGTGATATGCACTGGGTGAGGTTTGTGACAGAGG 873  
QY 789 CTGGGCACTCCCTCAAGGTTTTCATGCGGGAAGAGAGTACTTACAGGTATGACATT 848  
Db 874 TTGGCAACCCCATTTGAATGGCTTTATGACAGAGAGAGGAGTACTTTCAGTCCCTTCATT 933



Db 94 GAGCAATAACGGCAGAACTGGGAATCGAGAGCAACCAATGTCACCTACCAAGCCCA 153  
Qy 69 CCATGTGAGCAGGAATAAGAGAGGCAAGTGGTGGAAACAAGGGGTGGTTCGAGGATG 128  
Db 154 TCATGTCAGCAGGAAACAAGAGAGTCAAGTGGTGGGACCAAGAGTGGCTTCGTGGTGG 213  
Qy 129 TACCGTGTGCTAACAGAGTCTCTGGTGGCTGGAAACAACAGGATAAGTTTGGCCCTGGA 188  
Db 214 CACAGTTTGGCTACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCTTGGGA 273  
Qy 189 GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGCTCA 248  
Db 274 GGAGTACCTGTGTCTCATGTATTCATCTACACTCTGATGGTGCATATATCTGCA 333  
Qy 249 TGGCCTTAAACAGAAATCTCGGATCTCTCTGGGAGCAGAGAGAAATATCCCGCGAT 308  
Db 334 AGGTCTCAATAAATACTTGGCTTTAGTCTCTGAAAGCAGAGAGAGAAATGTTGCAAGCAT 393  
Qy 309 TGCTGAGGTGGCTAAGCTGTTTCTGATGCTGGTCTGGTCTGCATTAACAGCTTTATTTC 368  
Db 394 CGCAGAAGTTGCTAAACTGTTTGCAGATGCTGGCTTAGTGTGCATCAAGTTTCATATC 453  
Qy 369 TCCATTCCGAAAGGATCGTGAAGATGCCCGCAAAATACATGAATCAGCAGGCTGCCATT 428  
Db 454 ACCTTACACTCAGGATCGCAACATGCAAGGCATTTTCATGAAGTGCAGTTTACCGTT 513  
Qy 429 CTTTGAATATTTAGATGACCTCTAAATATTTTGTGAAGCAGAGACAGTAAAGCGCT 488  
Db 514 TTTTGAAGTATTTGTTGATGCTCTCTGCAATGTTTGTGAAGAGGATGTCAAGGACT 573  
Qy 489 CTATAAAGGGCCAGAGCTGGGAGATTAAAGGATTTACAGGATTTGATCTGATTATGA 548  
Db 574 CTACAAAAGCCGCGCAGGAGAAATTAAGGTTTCACTGGGATCGATTTGAAATATGA 633  
Qy 549 GAAACCTGAAACTCCTGAGGCTGTCTTAAACCAATTTGTCCACAGTGAAGTGTGT 608  
Db 634 AAAGCCAGAGGCCCTGAGTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693  
Qy 609 CCACAGGTAGTGAACCTTCTGCAAGAGCAGAACATTTGACCTTACTATATCAAGA 668  
Db 594 CCACAAAGTTGTGGAATCTTACAGGAGCGGATTTGTACCTGTGGATGATCTTATGA 753  
Qy 669 TATCCAGAACTCTTTGTCGGGAAACAACTTGCACAGCTCCGAGCTGAGGCTGAAC 728  
Db 754 AGTAAAGAACTATATGTCAGAGAAATAAATTCATTTGGCAAAAACAGATGCGGAAC 813  
Qy 729 TCTCCCTTCATATCAATTAAGTGAATCTCAAGTGGATCTCCAGTGGGTGAGGCAAG 788  
Db 814 ATTACAGCACTGAAAATTAATTAAGTGGATATGCAAGTGGGTGAGGCTTTTGGCAGAAG 873  
Qy 789 CTGGCCACCTCCCTCAAGGTTTCATGCGGGAGAGGAGTACTTACAGGTTATGCACTT 848  
Db 874 TTGGCAACCCCATTTGAATGGCTTTATGAGAGAGGAGGATCTTGCAGTGCCTTCAATTT 933  
Qy 849 TGACACCCCTGCTAGATGATGCGGTGATCAACATGAGCATCCCATTTCTACTGCCGCTC 908  
Db 934 TGATTGCTTCTGATGAGGTGCTATTAATTTCTAGTACCTATATCTTCTGACTGCGAC 993  
Qy 909 TGCAGAGGATAAGACAGCGGTGAGGAGGTGCAGCAAGTTTGTCTGTCACATGGTGGAGC 968  
Db 994 TCATGAAGATAAAGAGAGGCTGGACGCTGTACAGCATTTGCTCTGATGATGAGGCGCG 1053  
Qy 969 GAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAACAGAGAAAGAGAAAGCGTGTTC 1028  
Db 1054 CCGTGTGGCCATCTTCTGCAATCCAGAGTTTGTGAGCAGAGAAAGAGGAGCGCTGTGC 1113  
Qy 1029 CCGTGTGTGGGACAACTGTCAAAACACCCCATATCAAAATGGTGTGATGAAAGTGG 1088  
Db 1114 CAGACAGTGGGACGACATGCAAGAACACCCCTATATTAAGATGGTGTGAGCAAG 1173  
Qy 1089 GGACTGGCTGGTGGGAGACCTTCAAGTGTGCGAGAAATTAAGATGAATGATGGCT 1148  
Db 1174 AGATTGGCTGATTGGGAGGATCTTCAAGTCTTGGATCGAGTTTATTTGGAATGATGTCT 1233

Qy 1149 GGACCAATACCGTCTGACACCTCTGAGCTCAACACAGAAATGTAAGAAATGAATGCTGA 1208  
Db 1234 TGATCAGTATCGTCTTACTCTCTACTAGCTAAAGCAGAAATTTAAAGATATGAATGCTGA 1293  
Qy 1209 TGGCGTGTTCGATTCGATTCGAGTTGGCAATCCCTGTCCCAATGGCATGCCCTGTGATGCA 1268  
Db 1294 TGCTGTCTTTGATTTCACTTACCAACCCAGTGCACAATGGACATGCCCTGTAAATGCA 1353  
Qy 1269 GGACACCTGCCGCGAGGCTCTAGAGAGGGGTACAAAGCAGCCGGTCTCTCTACTACACCC 1328  
Db 1354 GGATACCAATGAAGCAACTTCTAGAGAGGGGTACCGGCGCCTGTCTCTCTCTCCACCC 1413  
Qy 1329 TCTGGCGGCTGGACCAAGGATGACGATGTCTCTTAGACTGGGGATGAAGCAGCAGC 1388  
Db 1414 TCTGGTGGCTGGCAAAAGGATGACGATGTCTTGTATGTGGGTATGAAGCAGCAGTGC 1473  
Qy 1389 GGCTGTCTCGAGGAAGGGTCTTGGATCCCAAGTCAACCAATTTTGGCCATCTTCCGTC 1448  
Db 1474 TGCAGTGTGGAGGAAGGATTTGAAATCTCTGAGACGACAGTGGTGGCCATCTTCCCATC 1533  
Qy 1449 TCCCATGTATATCTGCCCCCAGAGGTCAGTGGCAGTCCAGGATGCCCATCTCTGAACCAA 1508  
Db 1534 TCCCATGATGATGCTGACCAACTGAGTCCAGTGGCATTTGCAGACGAGGATGGTTC 1593  
Qy 1509 GGGTGCCTTATTTACATTTGGGGAGGGACCTTCAGGAATGCCCATCTCTGAACCAA 1568  
Db 1594 AGAGCCCAACTTTTACATTTGGGACGAGACCTCTGCTGGCATGCTCATCCAGAAACAGG 1653  
Qy 1569 GAAGGATCTGTATGAACCCACTCATGGGCAAGGCTTGAAGCATGGCCCTGGCCCTCAC 1628  
Db 1654 GAAGATCTTTATGAGCCAAAGTCAATGGTCCCAAGTGTGCTGACGATGGCCCTGGTTAAT 1713  
Qy 1629 CTCTGTGAAATCATCTCCATTCGAGTGGCTGCTTACAAAGAGCCAAAAGCCATGGA 1688  
Db 1714 CACTTTGAAATAGTTCCCTTTCAGTTGCACTTACAAAGAAAGAAAGCGTATGGA 1773  
Qy 1689 CTCTATATTCACAGCAGGACATGATTTGACTTTCATCTCAGGAATCGAATGAGAA 1748  
Db 1774 CTACTATGACTTGAACACCATGAAGACTTTTGAATTTATTTTCAGGAACACGAATGCGCA 1833  
Qy 1749 GCTGCCCGGAGAGAGGATCCCGAGATGGTTCATGGCCCGCCCAAGCATGGAAGGT 1808  
Db 1834 ACTTGTCTGAGAGGCGCAGAAACACCTGAGGTTTCATGGCTCCCAAGGCTTGGACGT 1893  
Qy 1809 CCGTACAGATTTATACAGTCCCTTGGAGAA 1838  
Db 1894 GCTGACAGAATACTACAAATCCTTGGAGAA 1923

## RESULT 5

US-09-966-262-116  
; Sequence 116, Application US/09966262  
; Publication No. US20030050461A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE OF INVENTION: P2004P1  
; CURRENT APPLICATION NUMBER: US/09/966,262  
; CURRENT FILING DATE: 2001-10-01  
; PRIOR APPLICATION NUMBER: US 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094



|    |      |  |      |
|----|------|--|------|
| Qy | 1449 | TCCATGTTATATGCTGCCCCACAGAGGTCCAGTGGCACTGCAGGTCCCCGATGATTGC   | 1508 |
|    |      |  |      |
| Db | 1534 | TCCCATGATGATGCTGAGCAACTGAGTTCAGTGGCATTTGAGAGCAGGATGGTTGC     | 1593 |
|    |      |  |      |
| Qy | 1509 | GGGTGCCAATTTCTACATTTGTGGGAGGGACCCTGCAGSAACTCCCCATCCTGAAACCAA | 1568 |
|    |      |  |      |
| Db | 1594 | AGGAGCCAACCTTTTACATTTTGGACGAGACCCTGCTGGCATGCTTCATCCAGAACAGG  | 1653 |
|    |      |  |      |
| Qy | 1569 | GAAGGATCTCTATGAACCCCACTCATGGGGCAAGGCTTTGAGCATGGCCCTGGCCTCAC  | 1628 |
|    |      |  |      |
| Db | 1654 | GAAGGATCTTTATGAGCCAAGTCATGTGCGCAAGTGCTCAGCATGGCCCTGGTTAAT    | 1713 |
|    |      |  |      |
| Qy | 1629 | CTCTGTGGAATCATTTCCATTCGAGTGGCTGCTACACAAAGCCAAAAAGCCATGGA     | 1688 |
|    |      |  |      |
| Db | 1714 | CACTTTGGAAATAGTTCCCTTTTCGAGTTGCAGCTTACACAGAAAAGAGCGTATGGA    | 1773 |
|    |      |  |      |
| Qy | 1689 | CTTCTATGATCCAGCAAGGCACAATGAGTTTGACTTCATCTCAGGAACCTCGAATAGGAA | 1748 |
|    |      |  |      |
| Db | 1774 | CTACTATGACTCTGAACACCATGAAGACTTTGAAATTTATTCAGGAACAGGAATGGCAA  | 1833 |
|    |      |  |      |
| Qy | 1749 | GCTCGCCGGGAAGGAGAGAAATCCCCAGATGCTTCATGGCCCCCAAGACATGGAAGGT   | 1808 |
|    |      |  |      |
| Db | 1834 | ACTTGTCTGAGAGGCCAGNAACCACTGNAGGTTTCATGGCTCCCAAGCCTTGGACCGT   | 1893 |
|    |      |  |      |
| Qy | 1809 | CCTGACAGATTTATACAGGTCCTCGAGAA                                | 1838 |
|    |      |  |      |
| Db | 1894 | GCTGACAGAACTACTACAAATCCTTTGGAGAA                             | 1923 |
|    |      |  |      |

## RESULT 6

```

RES001 TO
; Sequence 86, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095

```

|  |     |  |     |
|--|-----|--|-----|
| ; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,131<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,096<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,355<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,160<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,351<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/048,154<br>; PRIOR FILING DATE: 1997-05-30<br>; PRIOR APPLICATION NUMBER: US 60/054,804<br>; PRIOR FILING DATE: 1997-08-05<br>; PRIOR APPLICATION NUMBER: US 60/056,370<br>; PRIOR FILING DATE: 1997-08-19<br>; PRIOR APPLICATION NUMBER: US 60/060,862<br>; PRIOR FILING DATE: 1997-10-02<br>; NUMBER OF SEQ ID NOS: 343<br>; SOFTWARE: PatentIn Ver. 2.0<br>; SEQ ID NO 86<br>; LENGTH: 2527<br>; TYPE: DNA<br>; ORGANISM: Homo sapiens<br>; US-09-983-966-86 |     | Query Match 53.7%; Score 990; DB 9; Length 2527;<br>Best Local Similarity 71.3%; Pred. No. 0;<br>Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0; |     |
| QY   | 9   | GATCAAGAAGCAAAAGACGGAGAACACAGCAGCAAAATCCACCAATGTAGTCTATACGAGCCCA   | 68  |
| Db   | 94  | GAGCAATAACGGCGAGAAGCTGGGGAATGCAGAGAGCAACCAATGTCACTACCAAGCCCA   | 153 |
| QY   | 69  | CCATGTGAGCAGGAATTAAGAGAGGCAAGTGGTTGGAACAAGGGTGGGTTCCGAGGATG  | 128 |
| Db   | 154 | TCATGTTCAGCAGGAACAGAGAGGTGAGTGGTGGGACCAGAGAGTGGCTTCGTGGTTG   | 213 |
| QY   | 129 | TACCGTGTGGCTAACAGGTCTCTCTGCTGTGGAAAAACAACAGATAAGTTTTGCCCTGGA   | 188 |
| Db   | 214 | CACAGTTTGGCTAACAGGTTGTCCTGGAGCGGGAAGACACTACTGTGAGCATGGCCTTGA   | 273 |
| QY   | 189 | GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGGTCA  | 248 |
| Db   | 274 | GGAGTACCTGGTTGTTCATGGTATTCCATGCTACACTCTGGTGGTGACAATATTCGTCA  | 333 |
| QY   | 249 | TGGCTTTAACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATCCGCCGAT   | 308 |
| Db   | 334 | AGGTCTCAATAAATCTTGGCTTTAGTCTCTCTCTGGAAGACAGAGAGAAATGTTCCAGCAT  | 393 |
| QY   | 309 | TGCTGAGGTGGCTAAGCTGTTTGTGTGATGCTGGTCTGGTCTGCATTACCAAGCTTTATTTC   | 368 |
| Db   | 394 | CGCAGAAGTTGCTAAACTGTTTCAGATGCTGGCTTAGTGTGCATCACAGTTTCATATC   | 453 |
| QY   | 369 | TCCATTCCGAAGGATCGTGAGATGCCCGAAAAATACATGAATCAGCAGGGCTGCCATT   | 428 |
| Db   | 454 | ACCTTACACTCAGGATCGCAACAATGCAAGCGAAATTCATGAAGTGCAAGTTTACCGTT  | 513 |
| QY   | 429 | CTTTGAAATATTGTAGATCACTCTFAAATATTTTGTGAAGCAGAGACGCTAAAAAGCCCT   | 488 |
| Db   | 514 | TTTTGAAGTATTGTTGATGCTCTCTGTCATGTTTGTGAACACAGGGATGTCAAAGGACT  | 573 |
| QY   | 489 | CTATAAAGGGCCAGAGCTGGGAGATTAAGAGATTTACAGGTATTGATTCTGATTATGA   | 548 |
| Db   | 574 | CTCAAAAAAGCCGGCAGGAGAAATTAAGGTTTCACTGGGATCGATTCTCAATATGA   | 633 |
| QY   | 549 | GAACCTGAACTCCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAGTGACTGTGT  | 608 |
| Db   | 634 | AAAGCAGAGGCCCTTGAGTTGTGCTGAAACACAGACTCCTGTGTGTTAAATGACTGTGT  | 693 |
| QY   | 609 | CCACCGGTAGTGGAACTTCTGCAAGACGAGAAACATTTGACCCCTATACTATAATCAAGA   | 668 |



Db 694 CCAGCAAGTGTGGAACTTCTACAGGAACGGGATATTGTACTGTGGATGCATCTTATCA 753  
QY 669 TATCCAGCAACTCTTTGTGCGGGAACAAACCTTCACCAAGTCCAGCTGAGGCTGAAC 728  
Db 754 AGTAAAGCACTATATGTCCAGAAATAAATCTTATTTGGCAAAACAGATGCGGAAC 813  
QY 729 TCTCCCTTCATTTATCAATTAAGCTGATCTCCAGTGGGTCCAGGTTTGTAGCGAAG 788  
Db 814 ATTACAGCACTGAAATTAATAAAGTGGATGCGTGGGTGCGAGTTTGTGGCAAGG 873  
QY 789 CTGGGCCACTCCCTCAAGGTTTCATCGGGAAGAGGATCTTACAGGTTATGCACTT 848  
Db 874 TTGGCAACCCATTAATGAGCTTTATGAGAGAGGAGTACTTGCAGTGCCTCATTT 933  
QY 849 TGACACCTGTAGATGCGGTGATCAACATGAGCATGCCCATTTGACTGCCGCTC 908  
Db 934 TGATTGTCTCTGGATGGAGGTGCTAATTAAGTGTGCTGATGATGCTGCTGACGCGAC 993  
QY 909 TCCAGAGTAAAGACACCGCTGGGAAGGTGAGCAAGTTTGTCTGGGACATGTTGGAGG 968  
Db 994 TCATGAAGATAAAGAGAGGCTGGACGGCTGTACAGCATTTGCTCTGATGATGAGGCGG 1053  
QY 969 GAGGTAGCTATCTTACGAGCGCTGAATTTCTATGACACAGAAAGAGGAGCGTGTTC 1028  
Db 1054 CCGTGTGCCATCTTCCGAATCCAGAGTTTTTTCAGCAGAGGAAGAGGCGCTGTGC 1113  
QY 1029 CGGTGTTTGGGGACAACATGTACAAAACACCCCATATCAAAATGGTGTGAAGTGG 1088  
Db 1114 CAGACAGTGGGAACGACATCAAGAACCCCTATATTAAGATGGTGTGGAACAAG 1173  
QY 1089 GAGTGGCTGTGTGGAGACCTTTCAGGTGCTGGAGAAATAAGATGAATGATGGGCT 1148  
Db 1174 AGATTGGCTGTATGGAGGAGATCTTCAAGTCTTGGATCGAGTTTATTGGAATGATGGTCT 1233  
QY 1149 GACCAATACCTGTGACACCTCTGGAGCTCAACAGAAATCTAAGAAATCAATGCTGA 1208  
Db 1234 TGATCAGTATCGTCTACTCTACTGAGCTAAGCGAAATTTAAGATATGAATGCTGA 1293  
QY 1209 TCGCGTGTGATTCAGTTCCAGTTGCGCAATCTGTCCACAAATGCCCATGCCCTGTGATGCA 1268  
Db 1294 TGCTGTCTTGTGATTTCAACTACGCAACCCAGTGCACATGACATGCCCTGTGTAATGCA 1353  
QY 1269 GACACCTGCGCAGCGCTCTAGAGAGGGCTACAGCACCCGGTCTCTCTACTACACCC 1328  
Db 1354 GATACCCATTAAGCAACTCTAGAGAGGGGTACCGGCGCTGTCCTCTCTCCACCC 1413  
QY 1329 TCTGGCGCTGGACCAAGATGAGATGTGCTCTAGACTGGCGGATGAAGCAGCAGC 1388  
Db 1414 TCTGGTGTGGACAAAGGATGAGATGTTCCTTTGATGTGGCGTATGAAGCAGCATGC 1473  
QY 1389 GGCTGTGCTGAGGAAGGGTCTCGATGCCAAGTCAACCATTTGTTGCCATCTTTCCGTC 1448  
Db 1474 TGCAGTGTGGAGGAAGGATGCTGAATCTCTGAGACGACAGTGGTGGCGCATCTTCCCATC 1533  
QY 1449 TCCCATGTTATGCTGCGCCACAGAGTCCAGTGGCAGTCCAGTCCCGATGATGTC 1508  
Db 1534 TCCCATGATGATGCTGCGCAACTGAGGTCCAGTGGCATGTCAGACGACGAGTGGTTC 1593  
QY 1509 GGTGCCAATTTCTACATTTGGGAGGAGCCCTGCAGGAATGCCCATCTCTGAAACCAA 1568  
Db 1594 AGGACCAACTTTTACATTTTGGACGAGACCCCTGCTGGCATGCTCATCAGAAACAGG 1653  
QY 1569 GAAGATCTGATGAACCATCAATGCGGGCAAGGTCTTGAGCATGCGCCCTGGCCTCAC 1628  
Db 1654 GAAAGATCTTTATGAGCAAGTCAATGTTGCCAAGTGTGACGATGTCGAGTGGTGGTAAAT 1713  
QY 1629 CTCGTGGAATCATTTCCATTCGAGTCCGCTGCTTACAAACAGCAACCAAGCCATGGA 1688  
Db 1714 CACTTGGAAATAGTTCCTTTCGAGTTGCAAGTGTACAAACAGAAAGAACGCTGATGGA 1773  
QY 1689 CTTCTATGATCCAGCAAGCAATGAGTTTGTGACTTCTCAGGAATCGAATGAGGAA 1748  
Db 1774 CTACTATGACTCTGACACCATGAGACTTTGAATTTATTTTCAGGAACCAAGATGCGCAA 1833

QY 1749 GCTGCCCGGGAAGAGAGAAATCCCCAGATGCTTTCATGGCCCCCAAGCATGGAAGCT 1808  
Db 1834 ACTTCTCGAGAAGCCAGAAACACCTGGAAGGTTTCATGGCTCCCAAGGCTTGACCGT 1893  
QY 1809 CCGTACAGATTAATACAGGTCCCTCGAGAA 1838  
Db 1894 GCTGACAGAACTACTACAAATCCTTGGAGAA 1923

## RESULT 7

US-09-983-966-116  
; Sequence 116, Application US/09983966  
; Publication No. US20030060619A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004PI  
; CURRENT APPLICATION NUMBER: US/09/983,966  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,154  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/054,804  
; PRIOR FILING DATE: 1997-08-05  
; PRIOR APPLICATION NUMBER: US 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: US 60/060,862  
; PRIOR FILING DATE: 1997-10-02  
; NUMBER OF SEQ ID NOS: 343



| Matches | Conservative | 0; Mismatches   | 525; Indels | 0; Gaps | 0; |
|---------|--------------|---|-------------|---------|----|
| Qy      | 9            | GATCAAGAACAAAAGACGGAGAACACGACGAGAAATCCACCAAATAGTCTTATCATGAGCCCA | 68          |         |    |
| Db      | 94           | GAGCAATAACCGCGAGAACTGGGGAATGCGAGAGAGCAACCAATGTCACCTACCAAGCCCA   | 153         |         |    |
| Qy      | 69           | CCATGTGACAGGGAATAAGAGAGGGCAAGTGGTTGGAACAAGGGGTGGGTTCCGAGATG     | 128         |         |    |
| Db      | 154          | TCATGTGCAGGAAACAGAGAGGTGAGGTGGTGGGACCAAGAGGTGGCTTTCGTGGTGTG     | 213         |         |    |
| Qy      | 129          | TACCGTGTGGCTAACAGGTCTCTCTGTGCTGGAAAAACAACGATAAGTTTTGCGCTCGGA    | 188         |         |    |
| Db      | 214          | CACAGTTTGGCTAACAGGCTTGCTGGAGCGGGAAGACTACTGTGAGCATGGGCTTTGGA     | 273         |         |    |
| Qy      | 189          | GGAGTACCTTTGCTCCCATGCCATFCCCTTGTTACTCCCTGGATGGGACAAATGTCGGTCA   | 248         |         |    |
| Db      | 274          | GGAGTACCTGGTTTGTCTAGTATTCATGCTACACCTGGSATGTCGACATAATTCGTCA      | 333         |         |    |
| Qy      | 249          | TGGCCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGAAATATCCGCGGAT        | 308         |         |    |
| Db      | 334          | AGGTCTCAATTAATAATCTTGGCTTTAGTCTCTGAAGACAGAGAAGAGAAATGTTTCGACGAT | 393         |         |    |
| Qy      | 309          | TGCTGAGGTGGCTAAAGCTGTTTGGCTGATGCTGGTCTGCTGCAATACCACGCTTTATTTC   | 368         |         |    |
| Db      | 394          | CGCAGAAGTTGCTAAACTGTTTGCAGATCTGGCTTAGTGTGCATCACAACTTTCATATC     | 453         |         |    |
| Qy      | 369          | TCCATTTCGCAAAAGGATCGTGAGAAATGCCCGCAAAATACATGAATCAGCAGGGCTGCCATT | 428         |         |    |
| Db      | 454          | ACCTTACACTCAGGATCGCAACATTCGAAGCAAAATTCATGAAGTGCAAGTTTACCGTT     | 513         |         |    |
| Qy      | 429          | CTTTGAAATATTGTAGATGCACCTCTAAATATTGTTGAAAGCAGAGACGTAAGAAGCCT     | 488         |         |    |
| Db      | 514          | TTTTGAAGTATTTGTTGATGCTCCTCTGCAATGTTTGTGAACAGAGGGATGTCAAAGGACT   | 573         |         |    |
| Qy      | 489          | CTATAAAGGGCCAGACTCGGGAGATTAAAGGATTTACAGGTATTGATTCCTGATTATGA     | 548         |         |    |
| Db      | 574          | CTCAAAAAAGCCCGGCGAGAGAAATTAAGGTTTCTCTGGGATCGATCTGSAATATGA       | 633         |         |    |
| Qy      | 549          | GAAACCTGAAACTCCTCAGGGGTGCTTAAACCAATTTGCCACAGTGAAGTGTGTGT        | 608         |         |    |
| Db      | 634          | AAAGCCAGAGGCCCTCAGTTGGTGTCTGCTCAAAACAGACTCCTCTGATGTAAATGACTGTGT | 693         |         |    |
| Qy      | 609          | CCACCAGGTAGTGAACCTCTGCAAGAGCAGAACATTTGACCCCTATACTATAATCAAGA     | 668         |         |    |
| Db      | 694          | CCAGCAAGTTGTGAACCTCTACAGGAACGGGATATTGTACCTGTGGATGCATCTTATGA     | 753         |         |    |
| Qy      | 669          | TATCCAGCAACTCTTTGTGCGGGAACAAACAACCTTGACCACGTCGAGCTGAGGTGAAC     | 728         |         |    |
| Db      | 754          | AGTAAAAAGACTATATCTGCCAGAAAAATAACTTTCTTTGGCAAAAAACAGATGGGGAAC    | 813         |         |    |
| Qy      | 729          | TCCTCCCTTCATTTATCAATTACTAAGCTGGATCTCCAGTGGCTCCAGGTTTTGAGCGAAG   | 788         |         |    |
| Db      | 814          | ATTACAGCACTGAAAATTAATAAAGTGGATATGCAGTGGGTGCAGGTTTTTGGCAGAAG     | 873         |         |    |
| Qy      | 789          | CTGGGCCACTCCCTCAAGGTTTTTCATCGCGGGAAGAGTACTTACAGGTTATGCACTT      | 848         |         |    |
| Db      | 874          | TTGGCAACCCCATTTGAATGCTTTATGAGAGAGAGGAGTACTTTCAGTCCCTTCATT       | 933         |         |    |
| Qy      | 849          | TGACACCCCTGCTAGATGATGGCTGATCAACATGAGCATCCCCATTTGACTGCCCGTCTC    | 908         |         |    |
| Db      | 934          | TGATTGTCTTCTGATGGAGGTGCATTAACCTTCTCAGTACCTATAGTTCTGACTGCGAC     | 993         |         |    |





|  |    |      |  |      |
|--|----|------|--|------|
|  | Qy | 1215 | GTTTGGCAATCCAGTTGGCGAAATCCTGTCCACAATGCGCATGCCCTGTGTGATGCAGACAC | 1274 |
|  | Dd | 515  | CTTTGGCAATTCAACTACGCACCCGAGTGCAATGGACATGCCCTGTTAATGCAGGATAC    | 456  |
|  | Qy | 1275 | CTGCCGAGGCTCCTAGAGAGGGCTACAGCACCCCGCTCTCTACTACACCCTCTGGG       | 1334 |
|  | Dd | 455  | CCATAAGCAACTTCTAGAGAGGGCTACCGCGCGCTGTGCTCTCTCTCTCCCTCTGGG      | 396  |
|  | Qy | 1335 | CGGCTTGACCAGAAGATGACCATGTGCCCTCTAGACTGGCGGATG-AAGCAGCACCGCGCTG | 1393 |
|  | Dd | 395  | TGCGTGGCAAAGATGACCATGTCTCTTGTGTGGCGTATGAAGACGATGCTCGAC         | 336  |
|  | Qy | 1394 | TGCTCGAGGAAGGGTCTTGATCCCCAACCAATTGTGGCATCTTTCCGCTCTCCCA        | 1453 |
|  | Dd | 335  | TGTTGGAGGAAGAGTTCTGAATCCTGAGACGACAGTGGTGGCCAATCTCCCATCTCCCA    | 276  |
|  | Qy | 1454 | TGTTATATGCTGGCCGACAGAGGTCAGTGGGCATGTCAGGTCGCGGATGATTGCGGGTG    | 1513 |
|  | Dd | 275  | TGATGTATGCTGACCAACTGAGGTCCAGTGGCAATTGCAGAGCAGGATGGTTGCAGGAG    | 216  |
|  | Qy | 1514 | CCAATTTCTACATTGTGGGGAGGACCTCGACGGAATGCCCACTCTGAAACCAAGNAG      | 1573 |
|  | Dd | 215  | CCAACTTTTACATTGTGGACGAGACCCCTCTGGCATGCCCTCATCCAGAAAAGGGAAG     | 156  |
|  | Qy | 1574 | ATCTGTATGAACCCACTCATGGGGGAAGGTCATTGAGCATGGGCCCTGGCCTCACCTCTG   | 1633 |
|  | Dd | 155  | ATCTTTATGAGCCAAGTCATGGTGCCAAAGTGTCTGACGATGGCCCTGGTTTATCACTT    | 96   |
|  | Qy | 1634 | TGGAATATCATTTCCGAGTGGCTCTACAAACAAAGCCAAAAAGCCATGGACTTCT        | 1693 |
|  | Dd | 95   | TGGAATATGTTCCCTTTCGAGTTGCAGCTTTACAAACAAGAAAAGACGCTATGGACTACT   | 36   |
|  | Qy | 1694 | ATGATCCAGCAAGGCACATGATTTGCATTCAAT                              | 1727 |
|  | Dd | 35   | ATGACTCTGAACACCATTGAAGACTTTGAATTTAT                            | 2    |

RESULT 12

RE-501-12  
US-09-796--692-8898  
/ Sequence 8898, Application US/09796692  
/ Publication No. US20020198362A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Gaiger, Alexander  
/ APPLICANT: Algate, Paul A.  
/ APPLICANT: Mannion, Jane  
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
/ TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
/ FILE REFERENCE: 2077.001200  
/ CURRENT APPLICATION NUMBER: US/09/796,692  
/ CURRENT FILING DATE: 2001-03-01  
/ PRIOR APPLICATION NUMBER: 60/186,126  
/ PRIOR FILING DATE: 2000-03-01  
/ PRIOR APPLICATION NUMBER: 60/190,479  
/ PRIOR FILING DATE: 2000-03-17  
/ PRIOR APPLICATION NUMBER: 60/200,545  
/ PRIOR FILING DATE: 2000-04-27  
/ PRIOR APPLICATION NUMBER: 60/200,303  
/ PRIOR FILING DATE: 2000-04-28  
/ PRIOR APPLICATION NUMBER: 60/200,779  
/ PRIOR FILING DATE: 2000-04-28  
/ PRIOR APPLICATION NUMBER: 60/200,999  
/ PRIOR FILING DATE: 2000-03-01  
/ PRIOR APPLICATION NUMBER: 60/202,084  
/ PRIOR FILING DATE: 2000-05-04  
/ PRIOR APPLICATION NUMBER: 60/206,201  
/ PRIOR FILING DATE: 2000-05-22  
/ PRIOR APPLICATION NUMBER: 60/218,950  
/ PRIOR FILING DATE: 2000-07-14  
/ PRIOR APPLICATION NUMBER: 60/222,903  
/ PRIOR FILING DATE: 2000-08-03  
/ PRIOR APPLICATION NUMBER: 60/223,416  
/ PRIOR FILING DATE: 2000-08-04  
/ PRIOR APPLICATION NUMBER: 60/223,378

```

: PRIOR FILING DATE: 2000-08-07
:
: NUMBER OF SEQ ID NOS: 9597
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8898
: LENGTH: 600
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
:
: NAME/KEY: unsure
: LOCATION: (572)
: OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8898

```

|                           |       |                    |        |             |
|---------------------------|-------|--------------------|--------|-------------|
| Query Match               | 13.2% | Score 243;         | DB.9;  | Length 600; |
| Best Local Similarity     | 75.2% | Pred. NO. 1.1e-71; |        |             |
| Matches 303; Conservative | 0;    | Mismatches 100;    | Indels |             |

|      |     |  |      |
|------|-----|--|------|
| 1436 | Qy  | CCATCTTTCCGTCCTCCCATGTTATATGCTGGCCCCACAGAGGTCAGTGGCACTGCAGGT   | 1495 |
|      |     |  |      |
| 1    | Db  | CCATCTTCCCATCTCCCATGATGATGCTGGAGCAACTGAGGTCCAGTGGCATTCGAGAG    | 60   |
|      |     |  |      |
|      |     |  |      |
| 1496 | -Qy | CCCGAGTATTCGGGTGCCAATTCTACATTCGGGAGGAGCCCTGCAGGAATGCCCC        | 1555 |
|      |     |  |      |
|      |     |  |      |
| 61   | Db  | CACGGATGGTTCAGAGAGCCAATTTTACATTTGGACGAGACCCCTGCTGGCATGCCCT     | 120  |
|      |     |  |      |
|      |     |  |      |
| 1556 | Qy  | ATCCTGAAACCAAGAAGGATCTCTATGAACCCACTCATGGGGCAGAGTCTTGAGCATGG    | 1615 |
|      |     |  |      |
|      |     |  |      |
| 121  | Db  | ATCCAGAACAGGAGGATCTTTATGAGCCAAGTCATGCTGCCAAGTGCTCAGCATGG       | 180  |
|      |     |  |      |
|      |     |  |      |
| 1616 | Qy  | CCCTGGCTCACCCTCTGTGGAAATCATTCCTCCGAGTGGCTGCTACACAAAGCCA        | 1675 |
|      |     |  |      |
| 181  | Db  | CCCTGGTTTAATCACTTTTGGAAATAGTTCCTTTTCGAGTTGCAGCTTACAACAAGAAA    | 240  |
|      |     |  |      |
|      |     |  |      |
| 1676 | Qy  | AAAAAGCCATGGACTTCTATGATCCACGACAGGCACAAATGAGTTTGACTTCATCTCAGGAA | 1735 |
|      |     |  |      |
|      |     |  |      |
| 241  | Db  | AGAAGCGTATGGACTACTATGACTCTGAACACCATGAAGACTTTGAAATTTATTTTCAGGA  | 300  |
|      |     |  |      |
|      |     |  |      |
| 1736 | Qy  | CTCGAATGAGGAAGCTCGCCCGGAGAGGAGAGAAATCCCCAGATGCTTCATGGCCCCCA    | 1795 |
|      |     |  |      |
|      |     |  |      |
| 301  | Db  | CAGGAATGCCAANACTTGCTCGAGAGAGGCCAGAAACCACTGAAGGTTTCATGGCTCCCA   | 360  |
|      |     |  |      |
|      |     |  |      |
| 1796 | Qy  | AAGCATGGAAGTCCTGACAGATATTATACAGGTCCCTGGAGAA                    | 1838 |
|      |     |  |      |
|      |     |  |      |
| 361  | Db  | AGCTTTGGACCGTGGCTGACAGAAATACTACAAATCTCTGGAGAA                  | 403  |
|      |     |  |      |
|      |     |  |      |

RESULT 13

RESULTS 13  
US-10-040-862-8898  
; Sequence 8898, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Galger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01

```
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8898
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (572)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-8898

Query Match      13.2%; Score 243; DB 9; Length 600;
Best Local Similarity 75.2%; Pred. No. 1.1e-71;
Matches 303; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1436 CCATCTTCCGTCCTCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCAGTGCAGT 1495
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCATCTTCCATCTCCCATGATGATGCTGGACCAACTGAGTCCAGTGGCAGTGCAGAG 60

QY 1496 CCGGATGATTCGGGTGCCAATTCTACATTTGGGGAGGACCTGCAGGAATGCCCC 1555
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CACGATGTTGGAGGAGCACTTTTACATTTGGGACGAGACCTGCTGGCATGCCCTC 120

QY 1556 ATCCGTAACCAAGAGGATCTGTATGAACCCACTCATGGGGGCAAGTCTTCAGCATGG 1615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ATCCAGAAACAGGAAGGATCTTTATGAGCAAGTCATGTCGCAAGTGGCTGACGATGG 180

QY 1616 CCCCTGGCCCTACCTCTGTGGAAATCATTCATTCGGAGTGGCTGCCTACACAAAGCCA 1675
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCCCTGGTTTATCACTTTGGAATAGTTTCCCTTTTCGAGTTGCAGCTTACACAAAGAAA 240

QY 1676 AAAAGCCATGACTTCTATGATCCAGCAAGGCACAAATCAGTTTGACTTCATCTCAGGAA 1735
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGAAGCGTATGACTACTATGACTCTGAACACATGAAGACTTTGAATTTATTCAGGAA 300

QY 1736 CTCGAATGAGGAAGCTCGCCGGGAGGAGAGAATCCCGAGATGGCTTCATGGCCCCCA 1795
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CACGAATGCGCAAACTTGCTCGAAGAGGCCAGAAACCACTGAAGGTTTCATGGCTCCCA 360

QY 1796 AGCATGGAAGTCTCGACAGATTTATACAGTCCCTCGAGAA 1838
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AGGCTTGGACCGTCTGACAGAAATACATAAAATCCTTTGGAGAA 403
```

```
RESULT 14
US-10-198-846-5659
; Sequence 5659, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
```

```
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 620, 658, 687, 703, 727, 754, 764, 765, 783, 792
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5659

Query Match      12.7%; Score 233.4; DB 9; Length 803;
Best Local Similarity 99.6%; Pred. No. 2.4e-68;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1475 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 1534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 186

QY 1535 GGGACCTCGCAGGATGCCCATCTCTGAAACCAAGAGGATCTGTATGAACCCACATCATG 1594
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 GGGACCTCGCAGGATGCCCATCTCTGAAACCAAGAGGATCTGTATGAACCCACATCATG 246

QY 1595 GGGGCAAGTCTTGACATGGCCCTGGCTCACCTCTGTGGAAATCATTCATTCGCGAG 1654
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 GGGGCAAGTCTTGACATGGCCCTGGCTCACCTCTGTGTGGAAATCATTCATTCGCGAG 306

QY 1655 TGCTGCCTACACAAAGCAAAAGCCATGACTTCTATGATCCAGCAAGCA 1709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 TGCTGCCTACACAAAGCAAAAGCCATGACTTCTATGATCCAGCAAGCA 361
```

```
RESULT 15
US-10-198-846-12859
; Sequence 12859, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12859
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,
; LOCATION: 18, 19, 824, 825, 826
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12859

Query Match      12.7%; Score 233.4; DB 9; Length 826;
Best Local Similarity 99.6%; Pred. No. 2.5e-68;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1475 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 1534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 430
```

Qy 1535 GGGACCCCTGCGGAATGCCCCATCCTGAAACCAAGAGGATCTGTATGAACCCACTCATG 1594  
 Db 431 GGGACCCCTGCGGAATGCCCCATCCTGAAACCAAGAGGATCTGTATGAACCCACTCATG 490  
 Qy 1595 GGGGCAAGGTCTTGAGCATGGCCCTGGCCCTCACCTCTGTGGAAATCATTCATTCGGAG 1654  
 Db 491 GGGGCAAGGTCTTGAGCATGGCCCTGGCCCTCACCTCTGTGGAAATCATTCATTCGGAG 550  
 Qy 1655 TGGCTGCCTACAAACCAAGCCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCA 1709  
 Db 551 TGGCTGCCTACAAACCAAGCCCAAAAGCCATGGACTTCTATGATCCAGCAAGGTA 605

Search completed: June 4, 2003, 19:09:39  
 Job time : 298 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 15:10:31 ; Search time 2709 Seconds  
(without alignments)  
11030.149 Million cell updates/sec

Title: US-09-898-165B-9  
Perfect score: 1845  
Sequence: 1 atctcgggcatcaagaagca.....ggctccctggagaagaactaa 1845

Scoring table: IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estim.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gss.\*
- 18: em\_gss\_hum.\*
- 19: em\_gss\_inv.\*
- 20: em\_gss\_pln.\*
- 21: em\_gss\_vrt.\*
- 22: em\_gss\_fun.\*
- 23: em\_gss\_mam.\*
- 24: em\_gss\_mus.\*
- 25: em\_gss\_Other.\*
- 26: em\_gss\_pro.\*
- 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 907   | 49.2        | 956    | 14 | BQ957456    |
| 2          | 905.6 | 49.1        | 1766   | 11 | BC032513    |
| 3          | 866.2 | 46.9        | 951    | 9  | AL540583    |
| 4          | 854.6 | 46.3        | 886    | 9  | AL550460    |
| 5          | 853.6 | 46.3        | 915    | 14 | BQ924373    |
| 6          | 777.6 | 42.1        | 872    | 14 | BQ220909    |

|    |       |      |      |    |          |
|----|-------|------|------|----|----------|
| 7  | 757.2 | 41.0 | 860  | 13 | BI831135 |
| 8  | 743.6 | 40.3 | 949  | 14 | BQ926758 |
| 9  | 734.8 | 39.8 | 1041 | 13 | BI829344 |
| 10 | 722.6 | 39.2 | 1054 | 14 | BQ942867 |
| 11 | 704.8 | 38.2 | 874  | 14 | BQ943498 |
| 12 | 672.4 | 36.4 | 917  | 14 | BQ935947 |
| 13 | 629   | 34.1 | 683  | 12 | BG484421 |
| 14 | 624.8 | 33.9 | 650  | 12 | BE786139 |
| 15 | 624   | 33.8 | 727  | 13 | BI103516 |
| 16 | 613.8 | 33.3 | 620  | 13 | BM263737 |
| 17 | 582.8 | 31.6 | 1048 | 14 | BM903600 |
| 18 | 580.4 | 31.5 | 873  | 14 | BQ221395 |
| 19 | 579.4 | 31.4 | 684  | 12 | BG438005 |
| 20 | 558.8 | 30.3 | 767  | 9  | AI648997 |
| 21 | 546.8 | 29.6 | 972  | 14 | BQ896549 |
| 22 | 528   | 28.6 | 640  | 12 | BG823104 |
| 23 | 517.8 | 28.1 | 959  | 9  | AL574314 |
| 24 | 513.2 | 27.8 | 906  | 12 | BF182876 |
| 25 | 512   | 27.8 | 624  | 10 | BE292722 |
| 26 | 508   | 27.5 | 944  | 13 | BM458625 |
| 27 | 503   | 27.3 | 946  | 12 | BG490797 |
| 28 | 488.8 | 26.5 | 492  | 14 | BM708017 |
| 29 | 476.8 | 25.8 | 490  | 14 | BQ776777 |
| 30 | 465   | 25.2 | 584  | 12 | BG723809 |
| 31 | 459   | 24.9 | 576  | 10 | BE311856 |
| 32 | 441.4 | 23.9 | 870  | 13 | BI101136 |
| 33 | 440.2 | 23.9 | 549  | 13 | BI338434 |
| 34 | 436.2 | 23.6 | 557  | 10 | AW571670 |
| 35 | 433.8 | 23.5 | 741  | 13 | BI855664 |
| 36 | 425.8 | 23.1 | 709  | 14 | BQ571984 |
| 37 | 425.2 | 23.0 | 886  | 14 | BQ652788 |
| 38 | 424.6 | 23.0 | 938  | 12 | BF788130 |
| 39 | 424.4 | 23.0 | 712  | 12 | BG778039 |
| 40 | 423.6 | 23.0 | 1062 | 12 | BG116418 |
| 41 | 419.2 | 22.7 | 979  | 13 | BI257912 |
| 42 | 417.6 | 22.6 | 711  | 13 | BI649417 |
| 43 | 415.6 | 22.5 | 804  | 13 | BI852968 |
| 44 | 412.8 | 22.4 | 715  | 13 | BI081933 |
| 45 | 407.2 | 22.1 | 916  | 9  | AL517203 |

ALIGNMENTS

RESULT 1  
BQ957456  
LOCUS BQ957456 956 bp mRNA linear EST 21-AUG-2002  
DEFINITION AGNCOURT\_8786574 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6371441  
5', mRNA sequence.  
ACCESSION BQ957456  
VERSION BQ957456.1 GI:22372934  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 956)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DPFP/Gazdar  
CDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LRCM2545 row: j column: 18  
High quality sequence stop: 633.  
Location/Qualifiers  
1. .956

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6371441"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT      262 a      220 c      258 g      214 t      2 others
ORIGIN
Query Match
Best Local Similarity 98.18; Pred. No. 2.4e-255; Length 956;
Matches 938; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

Qy 480 AAAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAAGGATTTACAGGTATTGATTC 539
Db 1 AAAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAAGGATTTACAGGTATTGATTC 60
Qy 540 TGATTATGAGAACTGAACTCTGAGCGTGTGCTTTAAACCAATTTGTCACAGTGAG 599
Db 61 TGATTATGAGAACTGAACTCTGAGCGTGTGCTTTAAACCAATTTGTCACAGTGAG 120
Qy 600 TGACTGTGTCACACAGGTAGTGAACCTCTGCAAGAGCAGCAACATTTACCCCTATATAT 659
Db 121 TGACTGTGTCACACAGGTAGTGAACCTCTGCAAGAGCAGCAACATTTACCCCTATATAT 180
Qy 660 AATCAAGATATCCACGAACCTCTTTGTCGGGAAACAACTTGACACGTCGCGAGCTGA 719
Db 181 AATCAAGATATCCACGAACCTCTTTGTCGGGAAACAACTTGACACGTCGCGAGCTGA 240
Qy 720 GGCTGAACCTCTCCCTTCATATCAATTAAGTCTGATGATCTCAGTGGGTCAGGTTT 779
Db 241 GGCTGAACCTCTCCCTTCATATCAATTAAGTCTGATGATCTCAGTGGGTCAGGTTT 300
Qy 780 GAGCGAAGGCTGGGCCACTCCCTCAAGAGTTTCATCGGGGAGAGGACTTTACAGGT 839
Db 301 GAGCGAAGGCTGGGCCACTCCCTCAAGAGTTTCATCGGGGAGAGGACTTTACAGGT 360
Qy 840 TATGCACCTTTGACACCCCTGTAGATGATGCGGTGATCAACATGAGCATCCCCATTGTACT 899
Db 361 TATGCACCTTTGACACCCCTGTAGATGATGCGGTGATCAACATGAGCATCCCCATTGTACT 420
Qy 900 GCCCGTCTCTGCAGAGGATAAGACACGCGCTGGAAGGTGCGACGAAGTTTCTCTGGCACA 959
Db 421 GCCCGTCTCTGCAGAGGATAAGACACGCGCTGGAAGGTGCGACGAAGTTTCTCTGGCACA 480
Qy 960 TGTGTGACGAGGATAGCTATCTTACGAGACGCTGAATCTTATCAACACAGAAAGAGGA 1019
Db 481 TGTGTGACGAGGATAGCTATCTTACGAGACGCTGAATCTTATCAACACAGAAAGAGGA 540
Qy 1020 ACCTGTTCCTGTTGGGGACACATGATCAAAACACCCCCATATCAAAATGGTGAT 1079
Db 541 ACCTGTTCCTGTTGGGGACACATGATCAAAACACCCCCATATCAAAATGGTGAT 600
Qy 1080 GGAAGTGGGACTGCTGCTGGTGAGACCTTCAGGTGCTGAGAAATAAGATGGAA 1139
Db 601 GGAAGTGGGACTGCTGCTGGTGAGACCTTCAGGTGCTGAGAAATAAGATGGAA 660
Qy 1140 TGATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAGAAATGTAAGAAAT 1199
Db 661 TGATGGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAGAAATGTAAGAAAT 720
Qy 1200 GAATGCTGATGCGGTGTTTGCATTCACATCCAGTTGCGCAATCTGTCACATGGCCATGCCCT 1259
Db 721 GAATGCTGATGCGGTGTTTGCATTCACATCCAGTTGCGCAATCTGTCACATGGCCATGCCCT 780

```

```

Qy 1260 GTTGATGCAGGACACCTGCCCGAGGCTCTTAGAGAGGGGCTTACAGCACCGGGTCTCTCT 1319
Db 781 GTTGATGCAGGACACCTGCCCGAGGCTCTTAGAGAGGGGCTTACAGCACCGGGTCTCTCT 840
Qy 1320 ACTACACCCCTCTGG-GGGCTGGACCAAGGATGAGATGTCCTCTAGACTGGCGGATGA 1378
Db 841 ACTACACCCCTCTGGGGGGCTGGACCAAGGATGAGATGTCCTCTAGACTGGCGGATGA 900
Qy 1379 AGCAGCACGGG-GCTGTGCTCGAGGAAGGGTCTCTGGATCCCAAGTCAACCATGCT 1433
Db 901 AGCAGCACGGCTGGGCTTCCTCAAGAAAGGGTCTCTGGATCCCAAGTCAACCATGCT 956

RESULT 2
BC032513 1766 bp mRNA linear HTC 27-JUN-2002
LOCUS Homo sapiens, Similar to 3'-phosphoadenosine 5'-phosphosulfate
DEFINITION Synthesis 1, clone IMAGE:5492659, mRNA.
ACCESSION BC032513.1
VERSION BC032513.1 GI:21619118
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1766)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 69 Row: e Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20127474
This clone has the following problem: frame shifted.

FEATURES
Location/Qualifiers
1..1766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5492659"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_67"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT 486 a 366 c 476 g 438 t
ORIGIN
Query Match 49.1%; Score 905.6; DB 11; Length 1766;
Best Local Similarity 71.4%; Pred. No. 8.4e-255;
Matches 1193; Conservative 0; Mismatches 479; Indels 0; Gaps 0;

```





Db 445 CTGCCATCTTTGAATATTTGTAGATGCACCTCTAAATATTTGTGAAGCAGACGTA 504  
QY 481 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTCCT 540  
Db 505 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTCCT 564  
QY 541 GATTATGAGAACTGAAACTCCTGAGCGTGTGCTTTAAACCAATTTGTCCACAGTGAGT 600  
Db 565 GATTATGAGAACTGAAACTCCTGAGCGTGTGCTTTAAACCAATTTGTCCACAGTGAGT 624  
QY 601 GACTGTGCCACAGTAGTGAACCTTTGCAAGAGCAGAACATTTGACCCCTATATACTATA 660  
Db 625 GACTGTGCCACAGTAGTGAACCTTTGCAAGAGCAGAACATTTGACCCCTATATACTATA 684  
QY 661 ATCAAGATATACACGAACCTTTGTGCGGAAACAACTTGACACAGCTCCGAGCTGAG 720  
Db 685 ATCAAGATATACACGAACCTTTGTGCGGAAACAACTTGACACAGCTCCGAGCTGAG 744  
QY 721 GCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGTGGTCCAGGTTTG 780  
Db 745 GCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGTGGTCCAGGTTTG 804  
QY 781 ACAGGAGGCTGGCCCACTCCCTCAAAGGTTTCATGCGGAGAGGAGTACTACAGTT 840  
Db 805 ACAGGAGGCTGGCCCACTCCCTCAAAGGTTTCATGCGGAGAGGAGTACTACAGTT 864  
QY 841 ATGCACCTTTGACACC 855  
Db 865 ATGCACCTTTGACACC 879

RESULT 5

BO924373 915 bp mRNA linear EST 20-AUG-2002  
LOCUS AGENCOURT\_8794733 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6374672  
DEFINITION 5', mRNA sequence.

ACCESSION BO924373

VERSION BO924373.1

KEYWORDS GI:22339404

SOURCE EST.

ORGANISM human.

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo;

JOURNAL NIH-MGC http://imgc.ncbi.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: DCTD/DTP/Gazdar

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM2554 row: a column: 09

High quality sequence stop: 630.

Location/Qualifiers

1..915

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6374672"

/clone\_lib="NIH\_MGC\_18"

/tissue\_type="large cell carcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOT7; Site:1: XhoI; Site:2:

ECORI; cDNA made by oligo-dT priming. Directionally cloned

into EORI/XhoI sites using the following 5' adaptor:

GCACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

BASE COUNT 252 a 210 c 245 g 207 t 1 others

ORIGIN

Query Match 46.3%; Score 853.6; DB 14; Length 915;  
Best Local Similarity 98.1%; Pred. No. 1.2e-239;  
Matches 895; Conservative 0; Mismatches 14; Indels 3; Caps 3;

QY 467 AAAGCAGACGCTAAAGGCCCTCTATAAAGGCCAGAGCTGGGGAGATTAAGGATTTA 526  
Db 1 AAAGCAGACGCTAAAGGCCCTCTATAAAGGCCAGAGCTGGGGAGATTAAGGATTTA 60  
QY 527 CAGGTATTGATCTGATTATGAGAACTCTGAGCGTGTGCTTTAAACCAATTTG 586  
Db 61 CAGGTATTGATCTGATTATGAGAACTCTGAGCGTGTGCTTTAAACCAATTTG 120  
QY 587 TGTCCACAGTAGTACTGTGTCCACAGGAGTGTGAACTTTGCAAGAGCAGCAATTTG 646  
Db 121 TGTCCACAGTAGTACTGTGTCCACAGGAGTGTGAACTTTGCAAGAGCAGCAATTTG 180  
QY 647 TACCCTATATACTAAAGATATCCAGCAACTCTTTGCGCGGAAACAACTTTGACC 706  
Db 181 TACCCTATATACTAAAGATATCCAGCAACTCTTTGCGCGGAAACAACTTTGACC 240  
QY 707 AGTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGT 766  
Db 241 AGTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGT 300  
QY 767 GGTCCAGGTTTGTAGCGAAGGCTGGGCCACTCCCTCAAAGGTTTTCATGCGGGAAGG 826  
Db 301 GGTCCAGGTTTGTAGCGAAGGCTGGGCCACTCCCTCAAAGGTTTTCATGCGGGAAGG 360  
QY 827 AGTACTTACAGGTTATGCACTTTGACACCTCTGTAGATGATGCGGTCAACATCAGCA 886  
Db 361 AGTACTTACAGGTTATGCACTTTGACACCTCTGTAGATGATGCGGTCAACATCAGCA 420  
QY 887 TCCCCATTGCTGCCCCGCTCTCTGCAGAGGATAAGACACGGCTGGAAGGTCGACCAAGT 946  
Db 421 TCCCCATTGCTGCCCCGCTCTCTGCAGAGGATAAGACACGGCTGGAAGGTCGACCAAGT 480  
QY 947 TTGCTCTGCGACATGCTGACGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAAC 1006  
Db 481 TTGCTCTGCGACATGCTGACGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAAC 540  
QY 1007 ACAGAAAGAGGAACGCTGTTCCCGTGTGTTGGGGGACACATCTACAAAACACCCCATTA 1066  
Db 541 ACAGAAAGAGGAACGCTGTTCCCGTGTGTTGGGGGACACATCTACAAAACACCCCATTA 600  
QY 1067 TCAAAATGCTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGA 1126  
Db 601 TCAAAATGCTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGA 660  
QY 1127 AAATAGATGGATGAT-GGGCTGGACCAATACCTCTGACACCTCTGAGAGCTCAACAG 1185  
Db 661 AAATAGATGGATGATGGGCTGGACCAATACCTCTGACACCTCTGAGAGCTCAACAG 720  
QY 1186 AAATGTAAGAAATGAATGCTGATGCGGTGTTTGGCAATTCAGTTCGCAATCTCTGTCCAC 1245  
Db 721 AAATGTAAGAAATGAATGCTGATGCGGTGTTTGGCAATTCAGTTCGCAATCTCTGTCCAC 780  
QY 1246 AATGGCCATGCCCTGTTGATGAGGACACCTGCGCAGGCTCTCTAGAGAGGGGCTACAAG 1305  
Db 781 AATGGCCCTGCCCTGTTGATGAGGACACCTGCGCAGGCTCTCTAGAGAGGGGCTACAAG 840  
QY 1306 CACCCGGT-CCTCTCTACTACACCTCTGCGCGGCTGG-ACCAGGATGAGGATGTCCTC 1363  
Db 841 CACCCGGTCTCTCTACTACACCTCTGCGCGGCTGGNACCAAGGATGACCATGTGCCCTC 900  
QY 1364 TAGACTGCGGGA 1375  
Db 901 TAGAATGGCCGA 912



(Invitrogen). Research Genetics tracking code 013. Note: this is a NIH\_MGC Library."

|                       |                 |   |           |             |  |
|-----------------------|-----------------|---|-----------|-------------|--|
| BASE COUNT            | 235 a           | 198 c   | 213 g     | 214 t       |  |
| ORIGIN                |                 |   |           |             |  |
| Query Match           | 41.0%           | Score 757.2;  | DB 13;    | Length 860; |  |
| Best Local Similarity | 96.1%           | Pred. No. 2.8e-211;                                       |           |             |  |
| Matches 830;          | Conservative 0; | Mismatches 28;  | Indels 6; | Gaps 5;     |  |
| Qy                    | 203             | CCCATGCCATCCCTTGTACTCCCTGGATGGGACAATGTCGTCATGGCCTTAACAGAA | 262       |             |  |
| Db                    | 1               | CCCATGCCATCCCTTGTACTCCCTGGATGGGACAATGTCGTCATGGCCTTAACAGAA | 60        |             |  |
| Qy                    | 263             | ATCTCGGATCTCTCTGGGACAGAGAGGAAAATATCCGCGGATTCCTGAGTGGCTA   | 322       |             |  |
| Db                    | 61              | ATCTCGGATCTCTCTGGGACAGAGAGGAAAATATCCGCGGATTCCTGAGTGGCTA   | 120       |             |  |
| Qy                    | 323             | AGCTGTTGCTGATGCTGCTGCTGCTGATACAGGTTTATTTCTCCATTCGCAAGG    | 382       |             |  |
| Db                    | 121             | AGCTGTTGCTGATGCTGCTGCTGCTGATACAGGTTTATTTCTCCATTCGCAAGG    | 180       |             |  |
| Qy                    | 383             | ATCTGAGATGCCGCAAAATACATGAATCAGCAGGCTGCCATTTCTTGAATATTG    | 442       |             |  |
| Db                    | 181             | ATCTGAGATGCCGCAAAATACATGAATCAGCAGGCTGCCATTTCTTGAATATTG    | 240       |             |  |
| Qy                    | 443             | TAGATGCACCTCTAAATATTTGTGAAGCAGAGACGTAAAGGCTCTATATAAGGGCCA | 502       |             |  |
| Db                    | 241             | TAGATGCACCTCTAAATATTTGTGAAGCAGAGACGTAAAGGCTCTATATAAGGGCCA | 300       |             |  |
| Qy                    | 503             | GAGCTGGGAGATTAAGGATTTACAGTATTTGATTTGATTTGATTTGATTTGATTTG  | 562       |             |  |
| Db                    | 301             | GAGCTGGGAGATTAAGGATTTACAGTATTTGATTTGATTTGATTTGATTTGATTTG  | 360       |             |  |
| Qy                    | 563             | CTGAGCTGTGCTTAAACCAATTTGTCACAGTGTGCTGCTGCTGCTGCTGCTGCTG   | 622       |             |  |
| Db                    | 361             | CTGAGCTGTGCTTAAACCAATTTGTCACAGTGTGCTGCTGCTGCTGCTGCTGCTG   | 419       |             |  |
| Qy                    | 623             | AACCTTCTGCAAGACGAGCAATTTACCTTATCTATCTATCTATCTATCTATCTAT   | 682       |             |  |
| Db                    | 420             | AACCTTCTGCAAGACGAGCAATTTACCTTATCTATCTATCTATCTATCTATCTAT   | 479       |             |  |
| Qy                    | 683             | TTGTGCCGGAACAACTTACACAGCTCCAGCTGAGGCTGAACTTCCCTTCATTAT    | 742       |             |  |
| Db                    | 480             | TTGTGCCGGAACAACTTACACAGCTCCAGCTGAGGCTGAACTTCCCTTCATTAT    | 539       |             |  |
| Qy                    | 743             | CAATTACTAAGCTGGATCTCCAGTGGGTCCAGTGGTGGGAGGCTGGGCTCCCTCC   | 802       |             |  |
| Db                    | 540             | CAATTACTAAGCTGGATCTCCAGTGGGTCCAGTGGTGGGAGGCTGGGCTCCCTCC   | 599       |             |  |
| Qy                    | 803             | TCAAAGGTTTCATGCGGGAGAGGAGTACTTTACAGGTTATGCACTTTGACACCTGTG | 862       |             |  |
| Db                    | 600             | TCAAAGGTTTCATGCGGGAGAGGAGTACTTTACAGGTTATGCACTTTGACACCTGTG | 659       |             |  |
| Qy                    | 863             | ATGATGCGCTGATCAACATGAGCATCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTG | 922       |             |  |
| Db                    | 660             | ATGATGCGCTGATCAACATGAGCATCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTG | 719       |             |  |
| Qy                    | 923             | CACGCTGGAAGGTTGACAGCAATTTGCTGCGCACATGTGGAGGAGGCT-AGCTATC  | 981       |             |  |
| Db                    | 720             | CACGCTGGAAGGTTGACAGCAATTTGCTGCGCACATGTGGAGGAGGCT-AGCTATC  | 778       |             |  |
| Qy                    | 982             | TTACGACAGCTGAATTTATGAA-CACAGAAAAAGAGCAACGCTTTCCCTGTTGGGG  | 1040      |             |  |
| Db                    | 779             | TTACGACAGCTGAATTTATGAA-CACAGAAAAAGAGCAACGCTTTCCCTGTTGGGG  | 836       |             |  |
| Qy                    | 1041            | GACACATGTACAAAACACCCCA                                    | 1064      |             |  |
| Db                    | 837             | GACACATGTGTTCAAAACACCCCA                                  | 860       |             |  |

RESULT 8

BQ926758

LOCUS

949 bp mRNA linear EST 20-AUG-2002

BQ926758

DEFINITION

AGENCOURT\_8795627 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:6429903

5', mRNA sequence.

ACCESSION BQ926758

VERSION BQ926758.1 GI:22341789

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DRP/Gazdar

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2618 row: n column: 16

High quality sequence stop: 587.

Location/Qualifiers

1. 949

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6429903"

/tissue\_type="large cell carcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site: 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

BASE COUNT 261 a 224 c 248 g 215 t 1 others

ORIGIN

Query Match 40.3%; Score 743.6; DB 14; Length 949;

Best Local Similarity 98.4%; Pred. No. 3e-207;

Matches 782; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

Qy 467 AAACGACAGAGCTAAAGGCTCTATATAAGGGCCAGAGCTGGGAGATTAAGGATTTA|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Db 1 AAACGACAGAGCTAAAGGCTCTATATAAGGGCCAGAGCTGGGAGATTAAGGATTTA | 60 |  |  |  |
| Qy 527 CAGGTATTGATTCTGATTATGAGAACTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG | 586 |  |  |  |
| Db 61 CAGGTATTGATTCTGATTATGAGAACTCTGAGGCTGCTGCTGCTGCTGCTGCTG | 120 |  |  |  |
| Qy 587 TGTCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 646 |  |  |  |
| Db 121 TGTCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 180 |  |  |  |
| Qy 647 TACCTTATCTATATCAAGATATCCAGCACTCTTTGTCGCGGAAACCAACTTGACC | 706 |  |  |  |
| Db 181 TACCTTATCTATATCAAGATATCCAGCACTCTTTGTCGCGGAAACCAACTTGACC | 240 |  |  |  |
| Qy 707 ACGTCCGAGCTGAGGCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGT | 766 |  |  |  |
| Db 241 ACGTCCGAGCTGAGGCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGT | 300 |  |  |  |
| Qy 767 GGGTCCAGGTTTGGAGCGAAGGCTGGGCCACTCCCTCAAGGTTTTCATGCGGGAGAAG | 826 |  |  |  |
| Db 301 GGGTCCAGGTTTGGAGCGAAGGCTGGGCCACTCCCTCAAGGTTTTCATGCGGGAGAAG | 360 |  |  |  |
| Qy 827 AGTACTTACAGGTTATGCACTTTGACACCCCTGCTAGATGATGGCTGATCAACATGACA | 886 |  |  |  |

Db 361 AGTACTACAGGTTATGCACCTTTGACACCCCTGCTAGATGATGCGGTGATCAACATGAGCA 420  
QY TCCCCATTGTACTGCCGCTCTCTGCAGAGGATAGACACGCTGGAGGGTGCAGCAAGT 946  
Db TTTTCTCTGTCACATGCTGGACGAGGAGTATCTTACGAGAGCTGGAAGGCTGAGAGGCTGAGCAAGT 480  
QY TTGTCTCTGTCACATGCTGGACGAGGAGTATCTTACGAGAGCTGGAAGGCTGAGAGGCTGAGCAAGT 1006  
Db TTTTCTCTGTCACATGCTGGACGAGGAGTATCTTACGAGAGCTGGAAGGCTGAGAGGCTGAGCAAGT 540  
QY ACAGAAAAGAGGAGCACTGTTCCCGTGTGGGGGACACATGATACAAAACACCCCCATA 1066  
Db ACAGAAAAGAGGAGCACTGTTCCCGTGTGGGGGACACATGATACAAAACACCCCCATA 600  
QY TCAAAATGTTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTCGAGA 1126  
Db TCCAAATGTTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTCGAGA 660  
QY AAATAAGATGGAATGATGGGCTGGGACCAATACCGTCTGACACCTCTGGAGCTCAAAACAGA 1186  
Db AAATAAGATGGAATGATGGGCTGGGACCAATACCGTCTGACACCTCTGGAGCTCAAAACAGA 720  
QY AATGTAAGAAAT-GAATGCTGATGCGTG-TTTGATTCAGTTCGCGCAATCC-TGTCC 1243  
Db AATGTAAGAAATGGAATGCTGATGCGTG-TTTGATTCAGTTCGCGCAATCC-TGTCC 780  
QY 1244 ACAATGGCCATGCC 1258  
Db 781 CCAATGCCCATGCC 795

BI829344 1041 bp mRNA linear EST 04-OCT-2001  
603079477F1 NIH\_MGC\_l19 Homo sapiens cDNA clone IMAGE:5171058 5',  
RNA sequence.  
ACCESSION BI829344  
VERSION BI829344.1 GI:15940894  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1041)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1425 row: j column: 19  
High quality sequence stop: 841.

## FEATURES

## Source

1. 1041  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5171058"  
/clone\_lib="NIH\_MGC\_l19"  
/tissue\_type="medulla"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."

BASE COUNT 269 a 242 c 284 g 246 t  
ORIGIN

Query Match 39.8%; Score 734.8; DB 13; Length 1041;  
Best Local Similarity 93.1%; Pred. No. 1.2e-204;  
Matches 858; Conservative 0; Mismatches 52; Indels 12; Gaps 8;

QY 1 ATCTCGGGATCAAGAACAAAGACGGAACACAGCAAAATCCACCAATGTAGTCTAT 60  
Db 57 ATCTCGGGATCAAGAACAAAGACGGAACACAGCAAAATCCACCAATGTAGTCTAT 116  
QY 61 CAGCGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAAACAAGGGTGGGTTT 120  
Db 117 CAGCGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAAACAAGGGTGGGTTT 176  
QY 121 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTCTGGAAACAAGTAAGTTT 180  
Db 177 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTCTGGAAACAAGTAAGTTT 236  
QY 181 GCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTGTTTACTCCCTGGATGGGACAAT 240  
Db 237 GCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTGTTTACTCCCTGGATGGGACAAT 296  
QY 241 GTCCTCATGGCCTTAACAGAAATCTCGGATTCTCTCTGGGACAGAGAGGAAATATC 300  
Db 297 GTCCTCATGGCCTTAACAGAAATCTCGGATTCTCTCTGGGACAGAGAGGAAATATC 356  
QY 301 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTCTGATGCTGCTGCTGCATACACG. 360  
Db 357 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTCTGATGCTGCTGCTGCATACACG 416  
QY 361 TTTATTCTCCATTCCAAAGGATCGTGAGAAATGCCCGCAAAATACATGAATCAGCAGG 420  
Db 417 TTTATTCTCCATTCCAAAGGATCGTGAGAAATGCCCGCAAAATACATGAATCAGCAGG 476  
QY 421 CTGCCATTCTTTGAAA-TATTGTAGATGCACCTCTAAATATTTTGTGAAAGCAGACGT 479  
Db 477 CTGCCATTCTTTGAAA-TATTGTAGATGCACCTCTAAATATTTTGTGAAAGCAGACGT 536  
QY 480 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATTC 539  
Db 537 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGTATTGATTC 596  
QY 540 TGATTATGAAACCTGAAACTCCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAG 599  
Db 597 TGATTATGAAACCTGAAACTCCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAG 656  
QY 600 TGACTGTGCCACCGTAGTGAACCTC-TGCAAGAGCAGAACATTTACCTATATACTA 658  
Db 657 TGACTGTGCCACCGTAGTGAACCTC-TTGAAGAGCAGAACATTTACCTATATACTA 716  
QY 659 TATC- AAAGATATCAGCAACTTTGTGCCGGAACAACTTGACCACTGCCAGCTCCGAGCT 717  
Db 717 TAATCAAAAGATATCAGCAACTTTGTGCCGGAACAACTTGACCACTGCCAGCTCCGAGCT 776  
QY 718 GAGGCTGAAACTCTCCCTTCATTATCAAT- -ACTAAGCTGGATCTCCAGTGGTCCAGG 775  
Db 777 GAGGCTGAAACTCTCCCTTCATTATTTACTAAGCTGGATCTCCAGTGGTCCAGGT 836  
QY 776 TTTTGAGCGAAGCTGGGCCACTC-CCCTCAAAAGTTTTCATCGGGGAGAGAGTA- --- 830  
Db 837 TTTTGAGCGAAGCTGGGCCACTCGCCCTCAAAAGGTTTACATGCGCGGAAGAGAGGAC 896  
QY 831 CTTACAGGTATGCACTTTGACACCTGCT-AGATGATGGCGTGA-TCAACATGAGCATC 888  
Db 897 TTTACAGGTATGCACTTTGACACCTGTTTAACTGAAGGGGTGAGACAACTTGAGCGTG 956  
QY 889 CCCATTGTACTGCCCTCTCTGTG 910  
Db 957 CCGGTGTGACTGCCCGCTATG 978









JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/BTP/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM9639 row: c column: 11  
High quality sequence stop: 648.

FEATURES  
source  
1..650  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3877186"  
/clone\_lib="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."

BASE COUNT 184 a 145 c 170 g 151 t

ORIGIN

Query Match 33.9%; Score 624.8; DB 12; Length 650;  
Best Local Similarity 99.4%; Pred. No. 2.3e-172;  
Matches 648; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 538 TCTGATTATGAGAACTGAACTCCTGAGCGTGTCTTAAACCAATTTGTCCACAGTG 597  
DB 1 TCTGATTATGAGAACTGAACTCCTGAGCGTGTCTTAAACCAATTTGTCCACAGTG 60

QY 598 AGTGACTGTGTCCACCACTGAGTGGAACTTCTGCAAGACGAGACATTTGACCTATACT 657  
DB 61 AGTGACTGTGTCCACCACTGAGTGGAACTTCTGCAAGACGAGACATTTGACCTATACT 120

QY 658 ATAATCAAGATATCCACGAACTTTGTGCGGGAACAACTTGACCACTCCGAGCT 717  
DB 121 ATAATCAAGATATCCACGAACTTTGTGCGGGAACAACTTGACCACTCCGAGCT 180

QY 718 GAGGCTGAACTCTCCCTTCAATATCAATTAAGCTGATCTCCAGTGGTCCAGGTT 777  
DB 181 GAGGCTGAACTCTCCCTTCAATATCAATTAAGCTGATCTCCAGTGGTCCAGGTT 240

QY 778 TTGAGCGAAGCTGGGCCACTCCCTCAAAAGTTTTCATGCGGGAAGAGTACTTACAG 837  
DB 241 TTGAGCGAAGCTGGGCCACTCCCTCAAAAGTTTTCATGCGGGAAGAGTACTTACAG 300

QY 838 GTTATGCACTTTGACACCTCTAGATGATGGCGTGATCAACATGAGCATCCCATTTGA 897  
DB 301 GTTATGCACTTTGACACCTCTAGATGATGGCGTGATCAACATGAGCATCCCATTTGA 360

QY 898 CTGCGCGTCTCTCAGAGGATAGACACGCTGGAAGGTGACAGAGTTTGTCTCGGCA 957  
DB 361 CTGCGCGTCTCTCAGAGGATAGACACGCTGGAAGGTGACAGAGTTTGTCTCGGCA 420

QY 958 CATGGTGGAGGGGTAGCTATCTTACGAGACGCTGAATTCATGAACACAGAAAGAG 1017  
DB 421 CATGGTGGAGGGGTAGCTATCTTACGAGACGCTGAATTCATGAACACAGAAAGAG 480

QY 1018 GAACGCTGTTCGCGTGTGGGGGACAACTGTACAAAACACCCCATATCAAAATGGTG 1077  
DB 481 GAACGCTGTTCGCGTGTGGGGGACAACTGTACAAAACACCCCATATCAAAATGGTG 540

QY 1078 ATGGAAGTGGGACTGGCTGTGGTGGAGACCTTTCAGGTGCTGGAGAAAATAAGATGG 1137  
DB 541 ATGGAAGTGGGACTGGCTGTGGTGGAGACCTTTCAGGTGCTGGAGAAAATAAGATGG 600

QY 1138 AATGATGGGCTGGAACAATACCGTCTGACACCTCTGGAGCTCAAAACAGAAAT 1189  
|||||

Db 601 AATGATGGGCTGGACCAATA-CGTCTGACA-CTCTGGAGCTCAACAGAAAT 650

## RESULT 15

BI103516  
LOCUS  
DEFINITION 602889057F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:5044119  
5', mRNA sequence.  
ACCESSION BI103516  
VERSION BI103516.1 GI:14554409  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
AUTHORS 1 (bases 1 to 797)  
TITLE NIH-MGC http://mgc.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1121 row: a column: 16  
High quality sequence stop: 779.

FEATURES  
source

1..797  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5044119"  
/clone\_lib="NCI\_CGAP\_Kid14"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 202 a 195 c 222 g 177 t

1 others

Query Match 33.8%; Score 624; DB 13; Length 797;  
Best Local Similarity 87.2%; Pred. No. 4.3e-172;  
Matches 695; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

QY 935 GGTGCAGCAAGTTTCTCTGGCACATGTTGGACGGAGGTAGCTATCTTACGAGACGCTG 994  
DB 1 GGTGCAGCAAAATTTGGCTTGTATGACGAGTTCGAGGTCGCTCTATTACAGGACCTG 60

QY 995 AATTCTATGAACACAGAAAGAGACGCTGTTCCCGTGTTCGGGGACAACTGTACAA 1054  
DB 61 AATTCTATGAGCATAGAAAGAGGAGCGCTGTTCTCGTGTGGGGAACACCCATGCAA 120

QY 1055 AACACCCCATATCAAAATGGTGTATGGAAGTGGGACTGGCTGGTGTGGTGGAGACCTTC 1114  
DB 121 AGCACCCCATATCAAAATGGTGTATGGAAGTGGGACTGGCTGTGTGGTGGAGACCTAC 180

QY 1115 AGGTGCTGGAGAAAATTAAGATGGAATGGGTGGGCTGGACCAATACCGTCTGACACCTCTGG 1174  
DB 181 AGGTGCTAGAGAAATTAAGTGGGACGATGGGCTGGACCAATACCGCTTACGCTCTGG 240

QY 1175 AGCTCAACAGAAATCTAAAGAAATGAATGCTGATCGGTGTGTGATTCACCTCCGCA 1234  
DB 241 AGCTCAACAGAAATGTAAGACATGAATGCTGATCCGTTGTTCATTCAGCTTCGCA 300

QY 1235 ATCTCTTCCCAATGGCCATGCCCTGTGATGACGAGACACCTGCCGACGCTCCTAGAGA 1294  
DB 301 ATCTCTTCCCAATGGTCTGATGCCCTCTGATGACGAGACACCCGCCGACGCTCCTAGAGA 360

Search completed: June 4, 2003, 17:28:35  
Job time : 2720 secs.

**THIS PAGE BLANK (11/27/10)**